

Om protein - protein search, using sw model									
Run on: May 23, 2005, 18:56:58 ; Search time 167 Seconds (without alignments)									
Title: US-10-081-816-12									
Perfect score: 2645									
Sequence: 1 MRPSGRKVVKGQGQNSGHS TITVGVITLPPPPIMKUKRGK 512									
Scoring table: BLOSUM62									
Gapop 10.0 , Gapext 0.5									
Searched: 2105692 seqs, 386760381 residues									
Total number of hits satisfying chosen parameters: 2105692									
Minimum DB seq length: 0									
Maximum DB seq length: 00000000									
Post-processing: Minimum Match 0%									
Listing first 45 summaries									
Database : A_Geneseq_16Dec04:*									
1: geneseq1980b:*									
2: geneseq21990b:*									
3: geneseq2000b:*									
4: geneseq2001b:*									
5: geneseq2002b:*									
6: geneseq2003b:*									
7: geneseq2003b:*									
8: geneseq2004b:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query	Match Length	DB	ID	Description			
1	2645	100.0	512	4	ABB70017	Abb70017 Drosophil	RESULT 1		
2	2445	100.0	512	5	ABB30492	Aae30492 Fruit fly	ABB70017	ID	ABB70017 standard; protein; 512 AA.
3	405	15.3	454	4	ABB68914	Abb68914 Drosophil	XX	XX	XX
4	390	14.7	364	4	ABB75197	Abb75197 Drosophil	DT	DT	26-MAR-2002 (first entry)
5	390	14.7	364	5	ABB30504	Aae30504 Fruit fly	DB	DB	Drosophila melanogaster polypeptide SEQ ID NO 36843.
6	169	6.4	33	5	ABB30548	Aae30548	KW	KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
7	160	6.0	433	4	ABB75232	Abb75232 Drosophil	KW	KW	Drosophila melanogaster.
8	135.5	5.1	369	5	ABB30509	Aae30509	OS	OS	Drosophila melanogaster.
9	132.5	5.0	736	5	ABB30497	Abb67525 Drosophil	XX	XX	W0200171042-A2.
10	132.5	5.0	1299	4	ABB70025	Aae30497 Fruit fly	PR	PR	23-MAR-2000; 2000US-019163P.
11	132.5	5.0	1299	4	ABB70025	Abb70025 Drosophil	XX	XX	11-JUL-2000; 2000US-00614150.
12	130.5	4.9	372	5	ABB30510	Aae30510	DR	DR	(PEKE) PR CORP NY.
13	130.5	4.9	372	5	ABB30510	Aae30510	DR	DR	Venter JC, Adams M, Li PW, Myers EW;
14	129	4.9	436	4	ABB68943	Abb68943 Drosophil	DR	DR	WPI; 2001-656860/75.
15	129	4.9	436	4	ABB30511	Aae30511	XX	XX	N-PSDB; ABL14120.
16	127	4.8	348	4	ABB75207	Abb75207 Drosophil	PT	PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
17	122	4.6	410	4	ABB71166	Abb71166 Drosophil	PT	PT	PT
18	122	4.6	5	ABB30491	Aae30491	Abb75193 Drosophil	XX	XX	Disclosure; SEQ ID NO 36843; 21pp + Sequence Listing; English.
19	122	4.6	414	4	ABB75193	Aae30511	PS	PS	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA Sequences (AB11176-ABU30511), expressed DNA sequences (AB10180-ABU30511) and the encoded proteins (AB25773-AB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
20	120	4.5	450	5	ABB30537	Aae30537 Fruit fly	CC	CC	ABB4771 Entesococ
21	119	4.5	477	4	ABB75227	Abb67527 Drosophil	CC	CC	Abb67792 Drosophil
22	118.5	4.5	2008	4	ABU04090	Aau04090 Human cad	CC	CC	Add51295 Fruit fly
23	118.5	4.5	2008	5	ABG61703	Abg61703 Human cad	CC	CC	Aae30556 Fruit fly
24	118.5	4.5	2008	6	ABR43636	Abr43636 Human C1A	CC	CC	Abb5217 Drosophil
25	118.5	4.5	2008	7	ADC79417	Adc79417 Human C1A	XX	XX	Abb67593 Photorhab

SQ	Sequence 512 AA:	PT
Query	Match 100.0%; Score 2645; DB 4; Length 512;	receptor for controlling a pest population in an area.
Best	Local Similarity 100.0%; Pred. No. 7.7e-302;	
Matches	512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 MRPSGEKVKVKGQGNSGHSLSGMANYYRICKGDAVFLNAKPLNSANAOQVLYGVRKSI 60	
Qy	61 GLAERLDADYEAPPDRKKSSTDSTASNPERKPSVYRNDPINDPFLNPLKPLNSANAOQVLYGVRKSI 60	
Db	61 GLAERLDADYEAPPDRKKSSTDSTASNPERKPSVYRNDPINDPFLNPLKPLNSANAOQVLYGVRKSI 60	
Qy	121 ARAKFENNSASFIYSVVFVLLACVYGVYANRIRHVRSLSGPFEAVAYFLVNILPI 180	
Qy	121 ARAKFENNSASFIYSVVFVLLACVYGVYANRIRHVRSLSGPFEAVAYFLVNILPI 180	
Db	181 MIPILWYEARAKIAKLFNDMDFEVLYQSGHSPLKPLKROKAVVIAVPLSLSVVI 240	
Qy	181 MIPILWYEARAKIAKLFNDMDFEVLYQSGHSPLKPLKROKAVVIAVPLSLSVVI 240	
Db	181 MIPILWYEARAKIAKLFNDMDFEVLYQSGHSPLKPLKROKAVVIAVPLSLSVVI 240	
Qy	241 THVTMSDLNINQVQVPCILDNTLTAMGAWWILCEAMSITAHLLAERFQALKHGPAM 300	
Db	241 THVTMSDLNINQVQVPCILDNTLTAMGAWWILCEAMSITAHLLAERFQALKHGPAM 300	
Qy	301 VADYRVLWLRISKLTRDTGNALCYTFVMSLYLFFITLSYGMQLSRSQFGKIDGLT 360	
Db	301 VADYRVLWLRISKLTRDTGNALCYTFVMSLYLFFITLSYGMQLSRSQFGKIDGLT 360	
Qy	361 ITALMWIGLLFYICDEAHYASVNRNTFQKLLMVEBLNNMSDAQTEINNPLRTEMPS 420	
Db	361 ITALMWIGLLFYICDEAHYASVNRNTFQKLLMVEBLNNMSDAQTEINNPLRTEMPS 420	
Qy	361 ITALMWIGLLFYICDEAHYASVNRNTFQKLLMVEBLNNMSDAQTEINNPLRTEMPS 420	
Db	361 ITALMWIGLLFYICDEAHYASVNRNTFQKLLMVEBLNNMSDAQTEINNPLRTEMPS 420	
Qy	421 TINCGRFPDVNRTLFKGLLTMVYLWLLQFOQISPTDKDSEGANNTVDFVNDSLD 480	
Db	421 TINCGRFPDVNRTLFKGLLTMVYLWLLQFOQISPTDKDSEGANNTVDFVNDSLD 480	
Qy	481 NDMSLMGASTLSITTVGTTLPPIMKLGRKG 512	
Db	481 NDMSLMGASTLSITTVGTTLPPIMKLGRKG 512	

SQ	Sequence 512 AA:	PT
Query	Match 100.0%; Score 2645; DB 5; Length 512;	useful for identifying a compound which specifically binds to the
Best	Local Similarity 100.0%; Pred. No. 7.7e-302;	receptor for controlling a pest population in an area.
Matches	512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 MRPSGEKVKVKGQGNSGHSLSGMANYYRICKGDAVFLNAKPLNSANAOQVLYGVRKSI 60	
Qy	61 GLAERLDADYEAPPDRKKSSTDSTASNPERKPSVYRNDPINDPFLNPLKPLNSANAOQVLYGVRKSI 60	
Db	61 GLAERLDADYEAPPDRKKSSTDSTASNPERKPSVYRNDPINDPFLNPLKPLNSANAOQVLYGVRKSI 60	
Qy	121 ARAKFENNSASFIYSVVFVLLACVYGVYANRIRHVRSLSGPFEAVAYFLVNILPI 180	
Qy	121 ARAKFENNSASFIYSVVFVLLACVYGVYANRIRHVRSLSGPFEAVAYFLVNILPI 180	
Db	181 MIPILWYEARAKIAKLFNDMDFEVLYQSGHSPLKPLKROKAVVIAVPLSLSVVI 240	
Qy	181 MIPILWYEARAKIAKLFNDMDFEVLYQSGHSPLKPLKROKAVVIAVPLSLSVVI 240	
Db	181 MIPILWYEARAKIAKLFNDMDFEVLYQSGHSPLKPLKROKAVVIAVPLSLSVVI 240	
Qy	241 THVTMSDLNINQVQVPCILDNTLTAMGAWWILCEAMSITAHLLAERFQALKHGPAM 300	
Db	241 THVTMSDLNINQVQVPCILDNTLTAMGAWWILCEAMSITAHLLAERFQALKHGPAM 300	
Qy	301 VADYRVLWLRISKLTRDTGNALCYTFVMSLYLFFITLSYGMQLSRSQFGKIDGLT 360	
Db	301 VADYRVLWLRISKLTRDTGNALCYTFVMSLYLFFITLSYGMQLSRSQFGKIDGLT 360	
Qy	361 ITALMWIGLLFYICDEAHYASVNRNTFQKLLMVEBLNNMSDAQTEINNPLRTEMPS 420	
Db	361 ITALMWIGLLFYICDEAHYASVNRNTFQKLLMVEBLNNMSDAQTEINNPLRTEMPS 420	
Qy	421 TINCGRFPDVNRTLFKGLLTMVYLWLLQFOQISPTDKDSEGANNTVDFVNDSLD 480	
Db	421 TINCGRFPDVNRTLFKGLLTMVYLWLLQFOQISPTDKDSEGANNTVDFVNDSLD 480	
Qy	481 NDMSLMGASTLSITTVGTTLPPIMKLGRKG 512	
Db	481 NDMSLMGASTLSITTVGTTLPPIMKLGRKG 512	

RESULT	2	PT
AAE30492	ID AAE30492 standard; protein; 512 AA.	useful for identifying a compound which specifically binds to the
XX		receptor for controlling a pest population in an area.
AC	AAE30492;	
XX		
AC	AAE30492;	
XX		
DT	24-FEB-2003 (first entry)	
XX		
DE	Fruit fly gustatory receptor protein, Gr63f1.	
XX		
KW	Fruit fly: gustatory receptor; Gr; odorant receptor; Or; pesticide; Gr63f1.	
XX		
OS	Drosophila melanogaster.	
XX		
PN	W0200268393-A2.	
XX		
PD	06-SEP-2002.	
XX		
PP	22-FEB-2002; 2002WO-US005414.	
XX		
PR	23-FEB-2001; 2001US-0271319P.	
XX		
PA	(UYCO) UNIV COLUMBIA NEW YORK.	
XX		
PI	Axel R, Scott K;	
XX		
DR	WPI; 2002-698668/75.	
XX		
PT	Novel nucleic acid encoding insect gustatory or odorant receptor protein	Drosophila melanogaster polypeptide SEQ ID NO 33234.

RESULT 5
 ID AAE30504
 AC AAE30504;
 DT 24-FEB-2003 (first entry)
 DE Fruit fly gustatory receptor protein, Gr21D1.
 KW Gr21D1.
 OS Drosophila melanogaster.
 XX WO200268593-A2.
 XX PD 06-SEP-2002.
 XX PR 23-FEB-2002; 2002WO-US005414.
 XX PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Axel R, Scott K;
 XX DR WPI; 2002-698668/75.
 XX PN AAE30548
 PT Novel nucleic acid encoding insect gustatory or odorant receptor protein useful for identifying a compound which specifically binds to the receptor for controlling a pest population in an area.
 XX PS Disclosure; Page 183-185; 264pp; English.
 XX The invention relates to an isolated nucleic acid encoding an insect gustatory receptor (Gr) or odorant receptor (Or) protein which comprises seven transmembrane domains and a C-terminal domain comprising consecutive amino acids. The invention is useful for identifying a compound which activates the insect receptor or inhibits the activity of the insect receptor. The purified insect receptor protein is embedded in a lipid bilayer. The invention is sprayed for combating ingestion of crops by pest insects, combating disease-carrying insects in an area and controlling a pest population in an area. The invention is useful for detecting the presence of insect gustatory or odorant receptor and for inhibiting the function of the receptor in humans or animals or in biological fluids isolated from them. The invention is also useful for identifying or isolating other insect receptors and for combating pest nuisances. The present sequence is fruit fly gustatory receptor (Gr) protein. Sequence 364 AA;

QY 165 CSLWYINCNAFGTASRALSDALQTTGKPKQLTYRHLWVLDLSHMQGRAYSNM 224
 327 VPMSLVLFILITLSIGLMSQ-LSEGGRIGKGLITITALWIGLGLVYICDEAHYAVN 385
 225 GMYCLVIFTTITIATGSISETIDHATKVEGLVIVFCMGLJIVICDEAHYASRKG 284
 QY 386 TNFOKKULMVELNWMSDAQTEINMFRATENPSTINCGFFDVNRTLFKGLLTMVY 445
 285 LDFOTKLUNINLTAVDAATQEVEMLVAINKNPPINLDGYANINRELITNISFMATY 344
 QY 446 LVVLLQFQIS 455
 DB 345 LVVLLQFQIT 354

RESULT 6
 ID AAE30548
 AC AAE30548;
 DT 24-FEB-2003 (first entry)
 DE Fruit fly gustatory receptor peptide, Gr63F1.
 KW Gr63F1.
 OS Drosophila melanogaster.
 XX WO200268593-A2.
 XX PD 06-SEP-2002.
 XX PR 23-FEB-2001; 2001US-0271319P.
 XX PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Axel R, Scott K;
 XX DR WPI; 2002-698668/75.
 XX PN AAE30548
 PT Novel nucleic acid encoding insect gustatory or odorant receptor protein useful for identifying a compound which specifically binds to the receptor for controlling a pest population in an area.
 XX PS Disclosure; Page 183-185; 264pp; English.
 XX The invention relates to an isolated nucleic acid encoding an insect gustatory receptor (Gr) or odorant receptor (Or) protein which comprises seven transmembrane domains and a C-terminal domain comprising consecutive amino acids. The invention is useful for identifying a compound which activates the insect receptor or inhibits the activity of the insect receptor. The purified insect receptor protein is embedded in a lipid bilayer. The invention is sprayed for combating ingestion of crops by pest insects, combating disease-carrying insects in an area and controlling a pest population in an area. The invention is useful for detecting the presence of insect gustatory or odorant receptor and for inhibiting the function of the receptor in humans or animals or in biological fluids isolated from them. The invention is also useful for identifying or isolating other insect receptors and for combating pest nuisances. The present sequence is fruit fly gustatory receptor (Gr) protein. Sequence 364 AA;

QY 110 ICVLPVVRHGP---ARAKFEMNSASPIYSVVFVLLACTVGIVANRRI-HIVRSLSGP 164
 1 MGVMPFHRNPBKNTLRTGEGWSKGQWMAFIYSCQTTWVLRERVKCPVTPDKRF 60
 QY 165 BEAVIAVFLVNLIPIMIPI-LWEARKIAKLFENDWDDEFLVYQISGHSIPLKLURQKA 223
 DB 61 DRAIVNVFISLFTNPLPASWRHGPQIAFKNNTVQYKFFKIGS----- 110
 QY 224 VYATIVPLPSVLS--VIVHTVMSDLINQVNPVCLDN-----LTAMLGAW 269
 DB 111 ---PVPFPNLPLTWSLQVFSWLLSTAINUSQ--YFLQDPFLWTFAYPIIAMLNCF 164
 QY 270 ---WPLICEAMSTAHLLAERFQAKLKHIGPAMVADYRVLTRSKLTROTGNALCYTF 326
 DB 165 CSLWYINCNAFGTASRALSDALQTTGKPKQLTYRHLWVLDLSHMQGRAYSNM 224
 QY 327 VPMSLVLFILITLSIGLMSQ-LSEGGRIGKGLITITALWIGLGLVYICDEAHYAVN 385
 DB 225 GMYCLVIFTTITIATGSISETIDHATKVEGLVIVFCMGLJIVICDEAHYASRKG 284
 QY 386 TNFOKKULMVELNWMSDAQTEINMFRATENPSTINCGFFDVNRTLFKGLLTMVY 445
 DB 285 LDFOTKLUNINLTAVDAATQEVEMLVAINKNPPINLDGYANINRELITNISFMATY 344
 QY 446 LVVLLQFQIS 455
 DB 345 LVVLLQFQIT 354

Query Match 14.7%; Score 390; DB 5; Length 364;
 Best Local Similarity 27.3%; Pred. No. 3; 6e-36; Mismatches 152; Indels 40; Gaps 9;

Matches 101; Conservative 77; Mismatches 152; Indels 40; Gaps 9;

CC sequence which activates the insect receptor or inhibits the activity of the insect receptor. The purified insect receptor protein is embedded in a lipid bilayer. The invention is sprayed for combating ingestion of crops by pest insects, combating disease-carrying insects in an area and controlling a pest population in an area. The invention is useful for

CC detecting the presence of insect gustatory or odorant receptor and for
 CC inhibiting the function of the receptor in humans or animal or in
 CC biological fluids isolated from them. The invention is also useful for
 CC identifying or isolating other insect receptors and for combating pest
 CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)
 CC peptide

SQ Sequence 33 AA;

Query Match 6.4%; Score 169; DB 5; Length 33;
 Best Local Similarity 100.0%; Pred. No. 8e-12; 0; Mismatches 0; Indels 0; Gaps 0;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 420 STINCGGGRDVNRLPKGLTMTTYVLLQF 452
 ID AAB75332; AC AAB75332;
 XX 1 STINCGGGFDVNRLPKGLTMTTYVLLQF 33

RESULT 7

AB75332
 ID AAB75332 standard; protein; 433 AA.

XX 03-APR-2001 (first entry)

DB Drosophila gustatory receptor GR98B.1 protein sequence.

XX Gustatory receptor; fruit fly; taste; pheromone; semiochemical;
 KW crop damage; pest control.

XX Drosophila melanogaster.

OS

XX WO2000077208-A2.

XX PD 21-DEC-2000.

XX PR 14-JUN-2000; 2000WO-US016211.

XX PR 14-JUN-1999; 99US-0138668P.

PR 10-FEB-2000; 2000US-0181704P.

XX PA (UYA) UNIV YALB.

XX PI Carlson RJ, Clyne PJ, Warr CG;

XX DR WPI; 2001-051873/07.

XX N-PSDB; AAF63771.

XX PT New isolated nucleic acid molecule encoding Drosophila Gustatory Receptor
 PT protein useful for e.g. identification of compounds which may be used for
 PT pest management.

XX PS Claim 12; Page 214-215; 227pp; English.

XX This invention relates to polynucleotide sequences AAF63732 - AAF63777
 CC which encode Drosophila gustatory receptor proteins represented by
 CC sequences AAB75193 - AAB7528. The invention includes methods for
 determining gustatory receptor ligands. Also included is a method for
 CC modulating the expression of the DNA encoding the receptors. The DNA and
 CC protein sequences may be used for the identification of compounds, e.g.
 CC pheromones and other semi-chemicals, which may be used for pest
 management. The DNA sequences may also be used for behavioural studies
 CC involving gustatory systems in various organisms. Also, the DNA sequences
 CC may also be used to track down gustatory receptor genes in insects that
 CC damage crops or transmit disease

XX Sequence 433 AA;

Query Match 6.0%; Score 160; DB 4; Length 433;
 Best Local Similarity 20.9%; Pred. No. 6.3e-09;
 Matches 91; Conservative 72; Mismatches 139; Indels 134; Gaps 18;

Qy 102 PINFLRLITIGVLPVIRGPARAKFEMNSASFISVVFVLLACVGYVANNR-IHIVS- 159
 ID 27 PLOPFTRL---HGRRGVILGYACVYLISIMVIVSCTYVNTLQDIDKHEAE 79

DB 160 ---LSSGP-EBAVIAVFLVNLTIPIMIPIILWEARIAKLFNDWDDEPEV-LYQISG- 212
 Qy 80 DSSKUMGKAYAERSWWVAMFWNDQNLIL---NFRRLARYDDIALEIDANNASGGF 133

DB 213 ---HSLPLKLR-QKAVVIAINL---PILSVLSWVTHVMSDLNINQV 253

Qy 134 VGORHWHWPRFLALSVGLWIVLVLGTPRFTLVALGELYHWNKVKLIEILMQL-KC 192

DB 254 VPCILDNLTAMIGAWMFTICEMASITAHLLAERFOKALKHGPAMVADYVILURSK 313

Qy 193 TERCV---FVLLIVELLRGHL-QOISYLE---GN 223

DB 314 LTRDUTGNALCYTEVFMSHLYLPPFTLTSYGLMSQSLSGFGIKDGLTTALWNLGFLFYI 373

DB 224 QSRUVQELC---VALKRNQOLLAGRIGWGLVNEVSLYF---TSLTILPLNLTILQI 275

Qy 374 CDEAHYASVW---VTRNQKLUMLVEMTWSNDAQTRBNMFLRA---TENN 418

DB 276 VNHALLKISYNPNBCCQYSKLVKEPKRNFTYKQVIFIGRVTGTCULLSINIFSLCLSYBFC 335

Qy 419 PSTIN---CGGPPDVNRLPFCGL 438

DB 336 IOTNSISIRVHQNYCULSAEDYLKQGLREYSLQMEHLKLIPTCGSUDINLKPGCGM 395

Qy 439 LTTMVVYVLLQFOI 454

DB 396 VVTLPGYVILVQFKI 411

RESULT 8

AAE330509
 ID AAE330509 standard; protein; 369 AA.

XX AAE330509;

XX DT 24-FEB-2003 (first entry)

XX DR Fruit fly gustatory receptor protein, Gr39D1.

XX PA Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;
 KW Gr39D1.

XX OS Drosophila melanogaster.

XX WO200268593-A2.

XX PD 06-SEP-2002.

XX PS 22-FEB-2002; 2002WO-US005414.

XX PR 23-FEB-2001; 2001US-0271319P.

XX PA (UYCO) UNIV COLUMBIA NEW YORK.

XX PI Axel R, Scott K;

XX DR WPI; 2002-698668/75.

XX Novel nucleic acid encoding insect gustatory or odorant receptor protein
 PT useful for identifying a compound which specifically binds to the
 PT receptor for controlling a pest population in an area.

XX Disclosure; Page 193-194; 264pp; English.

CC The invention relates to an isolated nucleic acid encoding an insect
 CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises
 CC seven transmembrane domains and a C-terminal domain comprising
 CC consecutive amino acids. The invention is useful for identifying a
 CC compound which activates the insect receptor or inhibits the activity of

CC the insect receptor. The purified insect receptor protein is embedded in
CC a lipid bilayer. The invention is sprayed for combating ingestion of
CC crops by pest insects, combating disease-carrying insects in an area and
CC controlling a pest population in an area. The invention is useful for
CC detecting the presence of insect gustatory or odorant receptor and for
CC inhibiting the function of the receptor in humans or animals or in
CC biological fluids isolated from them. The invention is also useful for
CC identifying or isolating other insect receptors and for combating pest
CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)
CC protein

SQ

Sequence 369 AA;

XX

(PEKE) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EW;
WPI; 2001-6556850/75.
N-PSDB; ABN1628.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signaling and cell-cell interactions.

Priority: SRC ID NO 29367. 21nm + Sequence listing: English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 100 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL16175), expressed DNA
CC sequences (ABL01840-ABL6175) and the encoded proteins (ABB5737-
CC ABB20721). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp://ftp.wipo.int/pub/published_pct_sequences
XX

RESULT 9
 ABB67525
 ID ABB67525 standard; protein; 736 AA.
 AC ABB67525;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE *Drosophila melanogaster* polypeptide SEQ ID NO 29367.
 XX
 KW *Drosophila*; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS *Drosophila melanogaster*.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PP 23-MAR-2001; 2001WO-US0009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-0014150.

ID	AAE3047	RESIDU 10
XX	AAE30497	standard; protein; 736 AA.
AC	AAE30497;	
DT	24-FEB-2003	(first entry)
XX	Fruit fly gustatory receptor protein, Gr98A1.	
DE	Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;	
KW	Gr98A1.	
XX		

OS Drosophila melanogaster.
 XX
 PN WO200268593-A2.
 XX
 PD 06-SEP-2002.
 XX PP 22-FEB-2002; 2002WO-US005414.
 XX PR 23-FEB-2001; 2001US-0271319P.
 XX PA (UYCO) UNTV COLUMBIA NEW YORK.
 XX PT Axel R., Scott K.
 XX PI
 XX DR WPI; 2002-698668/75.
 XX PT Novel nucleic acid encoding insect gustatory or odorant receptor protein useful for identifying a compound which specifically binds to the receptor for controlling a pest population in an area.
 XX PS Claim 3; Page 170-173; 264pp; English.
 XX CC The invention relates to an isolated nucleic acid encoding an insect gustatory receptor (Gr) or odorant receptor (Or) protein which comprises seven transmembrane domains and a C-terminal domain comprising consecutive amino acids. The invention is useful for identifying a compound which activates the insect receptor or inhibits the activity of the insect receptor. The purified insect receptor protein is embedded in a lipid bilayer. The invention is sprayed for combating ingestion of crops by pest insects, combating disease-carrying insects in an area and controlling a pest population in an area. The invention is useful for detecting the presence of insect gustatory or odorant receptor and for inhibiting the function of the receptor in humans or animals or in biological fluids isolated from them. The invention is also useful for identifying or isolating other insect receptors and for combating pest nuisances. The present sequence is fruit fly gustatory receptor (Gr) protein sequence 736 AA;
 XX SQ Score 132.5; DB 5; Length 736;
 Best Local Similarity 21.0%; Pred. No. 66; Mismatches 134; Indels 89; Gaps 16; Matches 77; Conservative 66;
 Oy 102 PIMWPLRIGVLPVVRHGPRAKEEMNSASPIYSVVFULLACTVGYVANNR-IHIVRS- 159
 Db 427 PIGPFRTRL-----HKGRRGIVLGVACYLISIISLWIVYCTANTIVALQDQHFAE 479
 Oy 160 ---LSGPPEBAVAYLEFLVNLNPIMIPILWYEARKIAKLENWDDDFEV-LYYQISG-- 212
 Db 480 DSSKVMGNTQKVLVAMFWNQINIL---NFRRLARIIDDIADLBIDLNNSASGV 533
 Oy 213 ---HSLPLKLR-QKAVYIIVL-----PILSVLSVITRTHMDSLNINOW 254
 Db 534 GQRHWRFRFRFLAISVGLVILVLLVGLTPRTFLVGLGPYIHWTKVLTBILIMQL-KCT 592
 Oy 255 PYCILDNLTAMLGAWWPUCRMSITAHLLRERQKALHGIGPAAWADYRVUMLSKL 314
 Db 593 EYCV-----FVLLYELGRHHL-QQISVLE-----GNO 623
 Oy 315 TRDGNALCYTFVMSLYLEFTITISIYGLMOSLSEGGFKIGDIGITALWNLIGLYIC 374
 Db 624 SRSQVQELC---VALKQNLLAGRIGWLVNSLYF--TUSLILQYNELTIQIV 675
 Oy 375 DABHYASVNR--TNPQKULLMVELNWMSDQATBIMFLRATENPSTINGGAFDVAR 432
 Db 676 NWALIKSVNPNBCCOITEDYLILKNGLREVSLOMHLKLJ-----FTCGSLFDINL 726
 Oy 433 TLFKGJ 439
 Db 727 KFRGCV 732

RESULT 11
 ABB70025
 ID ABB70025 standard; protein: 1299 AA.
 XX
 AC ABB70025;
 XX DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster polypeptide SEQ ID NO 36867.
 XX KW Drosophila; developmental biology; cell signaling; insecticide; pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US009231.
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX PA (PERKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI; 2001-656860/75.
 XX DR-N-PSDB; ABL14128.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
 XX PS Disclosure; SEQ ID NO 36867; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL1617-ABL3511), expressed DNA sequences (ABL01840-ABL1615) and the encoded proteins (ABB5737-ABB5721). The sequence for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1299 AA;
 Query Match 5.0%; Score 132.5; DB 4; Length 1299;
 Best Local Similarity 19.3%; Pred. No. 6.6e-05; Mismatches 159; Indels 105; Gaps 19; Matches 81; Conservative 75; Mismatches 159; Indels 105; Gaps 19;
 Oy 96 FYRNIDPIMWPLRIGVLPVVRHGPRAKEEMNSASPIYSVVFULLACTVGYVANN 152
 Db 484 FHRAVSNVPIQSQYGLVSVNRLADVDIRFWCSPRILYSLIGIANSBGAVIN 543
 Oy 153 RIHTVRSISGPFEREAVIYLFLVNLIPMIPILWYEARKIAKLFNDPDEVLYQISG 212
 Db 544 VIKTIN---FRTSSTLISIVCUBHL---FFWRLAQWPRIMRTFHGVQFLRIVY 596
 Oy 213 HSLPLKLRQKAVYIIVL-----SVWITVHMTSDLNIN----- 251
 Db 597 RFGSYBYRKIRRIVFVTIVMSALVEHCLLIGNSPHLSNMRTOCKINNTYFESIYKER 656
 Oy 252 ---QWVPCILNDLTAMLGAW-----FLICEMSTA-HLIA 285
 Db 657 PHLYMLPYTF-----WMLPILAEWQNTAYPRSRPTDCFCIMCTGIGLAARFHOLY 706
 Oy 286 RFGQALKHGTGAAMWADYRVMLRSLKJTR-DTGNALCYTFVPM-SLYLUPITIUS 340
 Db 707 RRIAVHVKRKPVWTFEVREHYLAIKRVHLLDAAAPVLLAARGNNNSFICP----- 760

QY 341 TYGLMSQLEGGEGKIDG---DTITALW-NIG-----LLFYICDEAHYASUNVRTNFO 389
 Db 761 -----QLFNSP-.KQIGVDIYVMLAFWYISLGPAVRTLITFVA---SSIN--DYE 804
 QY 390 KKLMLYELNWNMSDACTEINMFLRATNPSTINCQGFFDVR-TLKGLLMTMVTLW 449
 Db 805 RKIVTALRDVPSRAWSTEVRQPSBQGLNDTTALSGSGFFYIITRSVLVAMGTTITYELMI 864

RESULT 12
 AABY5208
 ID AABY5208 standard; protein; 372 AA.
 XX
 AC AABY5208;
 XX
 DT 03-APR-2001. (first entry)
 XX
 DE Drosophila gustatory receptor GR39D.2a protein sequence.
 XX
 KW Gustatory receptor; fruit fly; taste; pheromone; semiochemical;
 XX
 crop damage; pest control.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200077208-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 14-JUN-2000; 2000WO-US016211.
 XX
 PR 14-JUN-1999; 99US-0138668P.
 PR 10-FEB-2000; 2000US-0181704P.
 XX
 PA (UYVA) UNIV YALE.
 XX
 PI Carlson PJ, Clyne PJ, Warr CG;
 XX
 DR WPI; 2001-061873/07.
 XX
 DR N-PSDB; AAF63747.
 XX
 PT New isolated nucleic acid molecule encoding Drosophila Gustatory Receptor
 PT protein useful for e.g. identification of compounds which may be used for
 PT pest management.
 XX
 PS Claim 12; Page 129-130; 227pp; English.

CC This invention relates to polynucleotide sequences AAF63732 - AAF63777
 CC which encode Drosophila gustatory receptor proteins represented by
 CC sequences AABY5193 - AABY5238. The invention includes methods for
 CC determining gustatory receptor ligands. Also included is a method for
 CC modulating the expression of the DNA encoding the receptors. The DNA and
 CC protein sequences may be used for the identification of compounds, e.g.
 CC pheromones and other semi-chemicals, which may be used for pest
 management. The DNA sequences may also be used for behavioural studies
 CC involving gustatory systems in various organisms. Also, the DNA sequences
 CC may also be used to track down gustatory receptor genes in insects that
 CC damage crops or transmit diseases
 XX Sequence 372 AA;

QY 229 VLPILSVLUVVITHVTMSLDINQV-----VPICLDNLTAMIGAWPLICAMI 279
 Db 133 WILVCLYIYAMWTHFGIMWLTQISRLVLTIGFVRCILAN-----FGFTCYGMVV 185
 QY 280 TAHLABERFQALKH-----GPAAMWADY-RVLMWLRSLKLTDTGNALCYTFVF- 328
 Db 186 IILKLUQVQVKOLEHLVUSTTISAGVACRTHDBILLQORELLAVYGVILFRIQ 245
 QY 329 --MSLYLFIFITLISIYGLMSQLEGGEGKIDGILTIPALWNIGLFLYICDEAHYASUNRT 386
 AC 295 NCTAKM-LTKVPRTGFLDRMIEKFLKLRQKPLILTYGSEFALDKSTLFK-LFTAIIFTY 352
 XX
 DT 04-6 LVLLOFQ 453
 DB 353 MVLVQPK 360

RESULT 13
 AAE30510
 ID AAE30510 standard; protein; 372 AA.
 XX
 AC AAE30510;
 XX
 DT 24-FEB-2003 (first entry)
 XX
 DE Fruit fly gustatory receptor protein, Gr39D2a.
 XX
 KW Fruit fly; gustatory receptor; Gr; odorant receptor; or; pesticide;
 XX
 PA Gr39D2a.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200268593-A2.
 XX
 PD 06-SEP-2002.
 XX
 PR 22-FEB-2002; 2002WO-US005414.
 XX
 PR 23-FEB-2001; 2001US-0271319P.
 XX
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Axel R, Scott K;
 XX
 DR WPI; 2002-698668/75.
 XX
 PT Novel nucleic acid encoding insect gustatory or odorant receptor protein
 PT useful for identifying a compound which specifically binds to the
 PT receptor for controlling a pest population in an area.
 XX
 PS Disclosure; Page 194-196; 264pp; English.
 CC The invention relates to an isolated nucleic acid encoding an insect
 CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises
 CC seven transmembrane domains and a C-terminal domain comprising
 CC consecutive amino acids. The invention is useful for identifying a
 CC compound which activates the insect receptor or inhibits the activity of
 CC the insect receptor. The purified insect receptor protein is embedded in
 CC a lipid bilayer. The invention is sprayed for combating ingestion of
 CC crops by pest insects, combating disease-carrying insects in an area and
 CC controlling a pest population in an area. The invention is useful for
 CC detecting the presence of insect gustatory or odorant receptor and for
 CC inhibiting the function of the receptor in humans or animals or in
 CC biological fluids isolated from them. The invention is also useful for
 CC identifying or isolating other insect receptors and for combating pest
 CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)
 XX
 Sequence 372 AA;

Query Match 4.9%; Score 130.5; DB 5; Length 372; Best Local Similarity 21.7%; Pred. No. 1.5e-05; Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17; CC
 QY 129 S~~A~~PIYSV~~W~~PF~~N~~LA~~C~~Y~~G~~Y~~V~~Y~~A~~NR~~H~~-~~V~~R~~S~~AS~~G~~-~~P~~RE~~A~~VI-~~A~~Y~~F~~U- 174
 Db 39 S~~T~~AIQIVV~~V~~G~~F~~MA~~L~~GA~~L~~AS~~L~~Y~~M~~ET~~K~~OT~~G~~NT~~F~~DN~~A~~VL~~I~~TT~~S~~Y~~Q~~LA~~N~~W~~R~~Q 98
 QY 175 ---VN~~L~~P~~I~~M-~~I~~PI~~W~~Y~~A~~RK~~I~~K~~U~~F~~N~~D~~M~~D~~P~~E~~V~~Y~~Y~~O~~S~~GH~~S~~LP~~K~~Q~~R~~AV~~T~~I 228
 Db 99 Q~~K~~Q~~Q~~Q~~V~~N~~Q~~L~~Q~~RL~~Q~~SQ~~V~~ELL~~Q~~PE~~P~~Y~~A~~P~~Q~~---~~F~~R~~M~~Y~~Y~~---~~R~~I 132
 QY 229 V~~P~~I~~L~~S~~V~~S~~V~~W~~T~~H~~T~~U~~M~~S~~N~~I~~N~~Q~~V~~---~~V~~P~~C~~I~~D~~N~~U~~T~~A~~M~~G~~W~~A~~W~~L~~C~~A~~M~~S~~I 279
 Db 133 W~~I~~V~~C~~L~~T~~Y~~G~~AM~~W~~H~~F~~G~~I~~W~~N~~L~~T~~W~~M~~Q~~I~~S~~R~~V~~L~~T~~I~~G~~F~~V~~Y~~C~~T~~W~~N~~ 185
 QY 280 T~~A~~M~~I~~A~~E~~R~~F~~Q~~K~~A~~R~~H~~I~~---~~G~~P~~A~~M~~A~~Y~~R~~V~~W~~I~~R~~K~~U~~S~~K~~L~~T~~D~~G~~I~~N~~C~~I~~Y~~T~~U~~F~~- 328
 Db 186 I~~L~~K~~K~~L~~U~~Q~~V~~Q~~K~~O~~L~~H~~E~~L~~V~~S~~T~~T~~I~~S~~M~~A~~G~~V~~A~~Q~~C~~L~~R~~H~~E~~I~~L~~L~~G~~Q~~R~~E~~I~~Y~~T~~ 245
 QY 329 -~~M~~S~~L~~Y~~L~~F~~I~~I~~T~~S~~I~~Y~~G~~M~~S~~Q~~S~~E~~R~~G~~F~~G~~I~~K~~D~~I~~G~~T~~I~~T~~A~~W~~N~~I~~G~~L~~F~~I~~C~~E~~D~~A~~H~~Y~~A~~S~~V~~N~~T~~ 386
 Db 246 V~~M~~O~~C~~I~~L~~F~~I~~Y~~S~~N~~L~~---~~E~~G~~F~~H~~H~~S~~N~~-~~D~~L~~V~~I~~F~~W~~L~~A~~M~~P~~F~~L~~I~~P~~L~~V~~V~~D~~I~~ 294
 QY 387 N~~F~~O~~K~~K~~U~~M~~E~~L~~N~~M~~N~~S~~D~~A~~T~~E~~I~~N~~M~~P~~R~~A~~T~~E~~M~~N~~F~~I~~S~~T~~I~~N~~C~~G~~G~~F~~O~~U~~N~~R~~-~~T~~L~~F~~G~~I~~L~~T~~M~~W~~Y~~ 445
 Db 295 N~~I~~C~~T~~A~~M~~L~~-~~T~~K~~V~~P~~R~~G~~T~~G~~L~~R~~M~~I~~R~~K~~L~~Q~~R~~K~~I~~T~~A~~Y~~G~~F~~A~~L~~D~~K~~S~~T~~L~~F~~K~~-~~L~~F~~A~~I~~T~~Y~~ 352
 QY 446 L~~V~~V~~L~~Q~~O~~F~~O~~ 453
 Db 353 M~~N~~I~~V~~Q~~P~~K~~360~~

RESULT 14

ABB68743
 ID ABB68743 standard; protein: 436 AA.
 XX
 AC ABB68743;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 33021.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PB CORP NY
 XX
 Venter JC, Adams M, Li PWD, Myers EW;
 PT XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL12846.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
 PT
 XX
 PS Disclosure; SEQ ID NO 33021; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB101840-AB116175), expressed DNA sequences (ABL102021). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

Sequence 436 AA;

Query Match 4.9%; Score 129; DB 4; Length 436;

Best Local Similarity 21.2%; Pred. No. 2.9e-05; Matches 94; Conservative 70; Mismatches 172; Indels 108; Gaps 22;

QY 85 A~~S~~N~~P~~E~~P~~F~~E~~P-~~S~~V~~T~~R~~N~~D~~P~~I~~N~~P~~W~~L~~R~~I~~G~~C~~V~~L~~P~~-~~I~~V~~R~~H~~G~~P~~A~~R~~K~~F~~M~~N~~S~~A~~S~~P~~I~~Y~~V~~F~~V~~ 140
 Db 30 A~~Q~~G~~C~~L~~E~~F~~O~~D~~P~~I~~G~~F~~Y~~G~~A~~R~~P~~L~~C~~A~~Q~~F~~C~~I~~M~~P~~N~~I~~R~~R~~D~~P~~O~~V~~E~~K~~F~~Y~~R~~S~~I~~G~~A~~V~~T~~G~~I~~ 89
 QY 141 L-~~L~~-L~~A~~C~~Y~~G~~V~~A~~N~~R~~H~~I~~V~~R~~S~~L~~G~~P~~F~~E~~R~~A~~V~~I~~A~~L~~F~~V~~N~~I~~L~~P~~M~~I~~P~~I~~I~~W~~Y~~---~~A~~R~~K~~I~~A~~ 194
 Db 90 L~~G~~G~~M~~K~~T~~L~~G~~-~~A~~N-~~I~~L~~F~~E~~G~~L~~N~~A~~K~~N~~G~~I~~V~~G~~L~~V~~F~~L~~I~~---~~V~~G~~M~~W~~N~~I~~N~~V~~G~~A~~N~~S~~W~~ 136
 QY 195 K~~L~~R~~M~~D~~P~~E~~V~~Y---~~V~~O~~I~~S~~H~~L~~P~~K~~R~~Q~~K~~A~~V~~T~~A~~W~~V~~P~~L~~S~~V~~---~~N~~I~~C~~Y~~G~~W~~F~~ 235
 Db 137 H~~I~~M~~L~~P~~W~~S~~S~~V~~D~~I~~M~~L~~U~~F~~P~~P~~Y~~K~~G~~R~~S~~---~~I~~R~~S~~K~~V~~N~~V~~A~~S~~V~~V~~L~~A~~V~~G~~D~~M~~Y~~Y~~A~~S~~G~~C~~S~~S~~ 192
 QY 236 L~~S~~V~~V~~I~~T~~H~~V~~M~~S~~---~~N~~I~~N~~Q~~V~~P~~C~~I~~D~~N~~I~~T~~A~~M~~-~~I~~G~~A~~W~~---~~W~~P~~L~~I~~C~~E~~275~~
 Db 193 M~~H~~I~~Q~~C~~T~~H~~M~~S~~R~~T~~E~~G~~L~~K~~E~~K~~F~~S~~D~~I~~M~~T~~I~~M~~P~~---~~N~~F~~P~~M~~C~~Y~~G~~W~~F~~ 248
 QY 276 A~~M~~S~~I~~T~~A~~H~~I~~A~~E~~R~~F~~Q~~K~~A~~R~~H~~I~~---~~P~~A~~M~~Y~~A~~D~~V~~Y~~R~~I~~W~~R~~L~~R~~K~~S~~I~~T~~D~~T~~G~~N~~A~~C~~Y~~F~~V~~ 328
 Db 249 F~~I~~W~~T~~S~~I~~G~~A~~Q~~R~~F~~O~~F~~A~~A~~V~~G~~A~~L~~E~~R~~H~~P~~E~~A~~M~~W~~D~~I~~R~~R~~H~~I~~C~~E~~A~~---~~S~~L~~V~~E~~A~~S~~301~~

QY 329 M~~S~~I~~X~~L~~F~~R~~I~~T~~S~~Y~~G~~M~~S~~Q~~S~~E~~R~~G~~F~~G~~I~~K~~D~~I~~G~~T~~I~~T~~A~~W~~N~~I~~G~~L~~F~~I~~C~~E~~D~~A~~H~~Y~~A~~S~~V~~N~~T~~ 380
 Db 302 M~~S~~N~~T~~V~~F~~S~~C~~A~~N~~N~~V~~V~~I~~C~~N~~A~~I~~A~~T~~K~~R~~H~~P~~I~~N~~T~~V~~F~~W~~S~~L~~I~~F~~L~~A~~R~~T~~S~~L~~V~~E~~M~~A~~S~~K~~H~~D~~A 361
 QY 381 S~~V~~N~~R~~T~~N~~F~~O~~K~~K~~L~~N~~E~~L~~N~~M~~N~~S~~D~~Q~~-~~E~~I~~M~~R~~L~~A~~T~~E~~M~~N~~F~~I~~S~~T~~I~~N~~C~~G~~G~~F~~O~~U~~N~~R~~T~~L~~F~~K~~G~~ 437
 Db 362 S~~L~~---~~L~~-~~I~~P~~R~~S~~I~~L~~V~~P~~S~~G~~O~~T~~W~~O~~R~~F~~--~~A~~D~~O~~I~~T~~S~~E~~F~~G~~L~~G~~S~~T~~R~~F~~C~~L~~T~~R~~K~~S~~F~~G 410
 QY 438 L~~T~~T~~M~~V~~T~~V~~L~~V~~L~~Q~~O~~F~~O~~I~~S~~T~~D~~K~~G~~ 461
 Db 411 M~~L~~A~~T~~V~~I~~V~~E~~M~~L~~---~~Q~~I~~D~~A~~K~~S~~H~~K~~G~~ 432

RESULT 15

AAB30491
 ID AAB30491 standard; protein: 436 AA.

XX
 AC AAB30491;

XX
 DT 24-FEB-2003 (first entry)

XX
 DE Fruit fly gustatory receptor protein, Gr61D1.

XX
 KW Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide; Gr61D1.

XX
 OS Drosophila melanogaster.

XX
 PN WO200268593-A2.

XX
 PD 06-SEP-2002.

XX
 PP 22-FEB-2002; 2002WO-US005414.

XX
 PR 23-FEB-2001; 2001US-0271319P.

XX
 PA (UYCO) UNIV COLUMBIA NEW YORK.

Search completed: May 23, 2005, 19:22:13
Job time : 175 secs

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2005, 19:15:31 ; Search time 43 Seconds (without alignments)

Sequence: 888.845 Million cell updates/sec

Title: US-10-081-816-12

Perfect score: 2645

Sequence: 1 MRPSGERVVKHGQNSHS. ITIVNGTILPPPPIMKUKGRKG 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Parents_AA: * /iaa/5A_COMB.pep: *
1: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6C_PCTUS_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	109	4.1	106	4 US-09-270-767-12643
2	107	4.0	404	4 US-09-270-767-39176
3	107	4.0	404	4 US-09-270-767-49176
4	104.5	4.0	323	3 US-09-134-001C-5200
5	104	3.9	433	4 US-09-491-577-74
6	102	3.9	907	4 US-09-198-452A-306
7	102	3.9	239	4 US-09-438-185A-295
8	101	3.8	239	4 US-09-710-279-1696
9	101	3.8	395	4 US-09-491-577-98
10	100.5	3.8	518	3 US-09-134-001C-4744
11	99.5	3.8	1900	4 US-09-949-016-1076
12	99	3.7	4	US-09-538-092-54
13	98.5	3.7	98	3.7
14	98	3.7	360	1 US-09-489-039A-8212
15	98	3.7	360	1 US-08-672-109B-6
16	98	3.7	360	2 US-08-842-045-6
17	98	3.7	360	2 US-08-780-749A-1
18	98	3.7	360	3 US-08-629-335B-6
19	98	3.7	360	3 US-08-629-335B-6
20	98	3.7	360	3 US-08-629-335B-6
21	98	3.7	360	4 US-09-709-666-4
22	98	3.7	600	4 US-09-134-000C-5694
23	97.5	3.7	459	3 US-09-97-889-22
24	97.5	3.7	459	4 US-09-098-079-22
25	97.5	3.7	456	3 US-09-134-001C-5593
26	97	3.7	309	4 US-09-328-352-7637
27	97	3.7	391	1 US-07-816-283-2

Query Match 4.1%; Score 109; DB 4; Length 106; Best Local Similarity 45.1%; Pred. No. 0.0012; Mismatches 19; Indels 0; Gaps 0; Matches 23; Conservative 9; MisMatches 19; Indels 0; Gaps 0;

ALIGNMENTS

Result No.	Qy	405 OTEINNIFLARTEMNISTINGCGFPDWRTEFKGLUTMVTWVLUQFOIS 455
1	46 QKREVENLIVVATKNNPIMNLGYANINRELTINISFATMVTWVLUQFOIKT 96	RESULT 2
2	US-09-270-767-33959	US-09-270-767-33959
3	; Sequence 33959, Application US/09270767	; Sequence 33959, Application US/09270767
4	; Patent No. 6703491	; Patent No. 6703491
5	; GENERAL INFORMATION:	; GENERAL INFORMATION:
6	; APPLICANT: Homburger et al.	; APPLICANT: Homburger et al.
7	; FILE REFERENCE: Nucleic acids and proteins of <i>Drosophila melanogaster</i>	; FILE REFERENCE: Nucleic acids and proteins of <i>Drosophila melanogaster</i>
8	; CURRENT APPLICATION NUMBER: US/09/270,767	; CURRENT APPLICATION NUMBER: US/09/270,767
9	; CURRENT FILING DATE: 1999-03-17	; CURRENT FILING DATE: 1999-03-17
10	; NUMBER OF SEQ ID NOS: 62517	; NUMBER OF SEQ ID NOS: 62517
11	; SOFTWARE: PatentIn Ver. 2.0	; SOFTWARE: PatentIn Ver. 2.0
12	; SEQ ID NO 42643	; SEQ ID NO 42643
13	; LENGTH: 106	; LENGTH: 106
14	; TYPE: PRT	; TYPE: PRT
15	; ORGANISM: <i>Drosophila melanogaster</i>	; ORGANISM: <i>Drosophila melanogaster</i>
16	; SEQ ID NO 33959	; SEQ ID NO 33959
17	; LENGTH: 004	; LENGTH: 004
18	; TYPE: PRT	; TYPE: PRT
19	; ORGANISM: <i>Drosophila melanogaster</i>	; ORGANISM: <i>Drosophila melanogaster</i>
20	; SEQ ID NO 3593	; SEQ ID NO 3593
21	; LENGTH: 106	; LENGTH: 106
22	; TYPE: PRT	; TYPE: PRT
23	; ORGANISM: <i>Drosophila melanogaster</i>	; ORGANISM: <i>Drosophila melanogaster</i>
24	; SEQ ID NO 3593	; SEQ ID NO 3593
25	; LENGTH: 106	; LENGTH: 106
26	; TYPE: PRT	; TYPE: PRT
27	; ORGANISM: <i>Drosophila melanogaster</i>	; ORGANISM: <i>Drosophila melanogaster</i>

Query Match 4.0%; Score 107; DB 4; Length 404; Best Local Similarity 18.2%; Pred. No. 0.016; Matches 75; Conservative 85; MisMatches 155; Indels 96; Gaps 17; Gaps 17; Matches 75; Conservative 85; MisMatches 155; Indels 96; Gaps 17; Sequence 569, APPL Sequence 16, APPL Sequence 7, APPL Sequence 1, APPL Sequence 3, APPL Sequence 3105, APPL Sequence 2116, APPL Sequence 8, APPL Sequence 4395, APPL Sequence 4379, APPL Sequence 7912, APPL Sequence 17058, APPL Sequence 38, APPL Sequence 4, APPL Sequence 4, APPL Sequence 4276, APPL

Db 23 FMTVPGULLNRYVAGGRERFRFSKANLAFASLWIAFLSIVYG-----ROIKEYQ 72

QY 166 E-----AVIATLFL-----YVILPIMIPIIWYRARKIAKULFNDMDDEVIYVQISG 212

Db 73 EGQINLKDTATLYSMNITVAVINVSQMI-----SDHAKVULSKUPPFDTL----- 120

Db 213 HSLPLKURQKAVVIAVILPILSYLSWIT-----HVTMS-----DININ 251

Db 121 --KEFRUDRSLSLYTISIVLALKVTKAFLPLTEVAFPLQFQRRQHPEMSLWTLYRFLPLIS 178

Qy 252 QVVPYCLDNLNTAMLGAWMFL--ICAMSITAHIAERFOKA-----RHIGPAMWAD-- 303

Db 179 NPLFLNVCYFGAMVUVKETYLALNRLEAQEOLQEVNLQKDLKLYTKYTMQRCALABEL 238

Db 304 ----YVLMWLRSLKLTTRDTG----NACY----TWFVMSLUVLFFITLSIYGLMSOL 348

Qy 304 ----YVLMWLRSLKLTTRDTG----NACY----TWFVMSLUVLFFITLSIYGLMSOL 348

Db 239 DQLAYRVLIVHSGKVITPMSLMSLMSLICHIGITGFGSYLVAIDL-----IMGR 294

Qy 349 SBCFG--IKDIGHTITALWNIGLFLYICDDEAHYASVNRTNEOKKLAVELNNNNDA-- 404

Db 295 YDGLSLINLVFLSI-SLABITLTHLCNHLVAT-----RASAVILQEMNLOHADRY 347

Qy 405 QPRINMELURATEKPNSTINGGFFDVRNTRFLKGSLTTMVTYVLUVLOPOIS 455

Db 348 RQAVHGFITLVTYVQIKPLGLYEDMRLISNVFASAVASFLILVQADLS 398

RESULT 3

US-09-270-767-49176 Application US/09270767

; Sequence 49176, Application US/09270767

; Patient No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*

; FILE REFERENCE: 7326-094

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1999-03-17

; CURRENT FILING NUMBER: US 60/055,779

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 49176

; LENGTH: 404

; TYPE: PRT

; ORGANISM: *Drosophila melanogaster*

US-09-270-767-49176

Query Match 4.0%; Score 104.5; DB 3; Length 323;

Best Local Similarity 22.7%; 34; Pred. No. 0.02; Gaps 10;

Matches 58; Conservative 34; Mismatches 79; Index 85; Gaps 10;

Qy 271 FLCIAMSITAHIAERFOKA-----RVLWLRSLKLTTRDTG-----HACLYTPEVFM 329

Db 66 FIDNEBAYQHLSNSFPEPMNSPNHL---LGDDYGRDLSRSLVGSRAT-----LFVTL 116

Qy 330 SLYVFLITISIYGLMSOLSGF-----GKDIGITI-----TALWNIGL 369

Db 117 ITLFLFVVGVPGVGLIAGYKKGWIDTIPIMRIDIGLISIPEFVIMALASFFHPSLWNLV 176

Qy 370 IYICDABAHVAVSNVNTNFQKGLLMEVNNNNDAOTEINMELURATEKPNSTINGGFF 429

Db 177 ATI-----IKNN-----YTRVIGIUNTEMNOSYIOMQFN 210

Qy 430 VN--RTLFKGILTMTVYVLUVLUQFOISIPTDKGDSSEGANNITVWDVFMDSLNDMSLMG 487

Db 211 VSTVNILFKGLPKLKPULSIFVIM-----IVDP-----GKILY 243

Qy 488 ASTLSITVGTPLRPP 503

Db 244 ISSLSFLGIGAQPPSP 259

RESULT 5

US-09-491-577-74

; Sequence 74, Application US/09491577

; Patient No. 6610511

; GENERAL INFORMATION:

; APPLICANT: Yale University

; APPLICANT: Carlson, John R.

; APPLICANT: Kim, Hunyong

; APPLICANT: Clyne, Peter J.

; APPLICANT: Warr, Coral G.

; TITLE OF INVENTION: No. 6610511 Family of Odorant Receptor Genes in *Drosophila*

; FILE REFERENCE: 44574-5061-US

; CURRENT APPLICATION NUMBER: US/09/491,577

; CURRENT FILING DATE: 2000-01-25

; EARLIER APPLICATION NUMBER: US 60/117,132

; EARLIER FILING DATE: 1999-01-25

; NUMBER OF SEQ ID NOS: 112

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 74

; LENGTH: 363

RESULT 8
US-09-491-577-98
Query Match 3.8%; Score 101; DB 4; Length 395;
Best Local Similarity 18.9%; Pred. No. 0.066;
Matches 56; Conservative 58; Mismatches 112; Indels 70; Gap 12;
Db 408 VCLVRSYVAKTLKLNLSAQALQITRSQDWIKSMVTKQKROVFLFLAHLKHPSERHQT 467
; Sequence 1696, Application US/09710279
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PI3440TS
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1696
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1696

Query Match 3.8%; Score 101; DB 4; Length 239;
Best Local Similarity 23.7%; Pred. No. 0.03; Mismatches 61; Indels 82; Gaps 9;
Matches 53; Conservative 28; Mismatches 61; Indels 82; Gaps 9;
Db 303 DY-RVTLWRLSLKLTDRDGNALCYTFVFLMSLYVFFITLSIVGIMSQSEGII-----GIK 355
11 DYGDRDLSRLVGSRAT-----LFTFLTLLFTVVGVPLGILLAGYKKGMIDTIMRII 64
Query 356 DIGLTI-----TALWNGLLFLYCFDBAHYASTVNRTRNFQKKLMVNLWMN 401
65 DIGLSDPEFVNTIALASFFHDSSLWNVIAITI-----TRKWMN 101
Db 402 SDAQTEINMFLAREMPSTINCGGFDVN--RTUFGGLMTNTVWVLUQFQISIPTD 459
102 --YTRVTRGIVNTENMOSYQIMAOFFNVSTNLFLKHLPRVLPISIVIM-----149
Query 460 KGDSEGANNITVDFWMDSLDMDMSLNGASTIISTTWGTLPP 503
Db 150 -----IVDP-----GKIIIVISSIFLGAQPPSP 175
RESULT 9
US-09-491-577-98
; Sequence 98, Application US/09491577
; General Information:
; APPLICANT: Yale University
; APPLICANT: Carlson, John R.
; APPLICANT: Kim, Hunhyong
; APPLICANT: Clyne, Peter J.
; APPLICANT: Warr, Coral G.
; Title of Invention: No. 661051el Family of Odorant Receptor Genes in Drosophila
; File Reference: 44574-5061-US
; Current Application Number: US/09/491,577
; Current Filing Date: 2000-01-25
; Earlier Application Number: US 60/117,132
; Earlier Filing Date: 1999-01-25
; Number of SEQ ID NOS: 112
; Software: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 395
; TYPE: PRT

RESULT 10
US-09-134-001C-4744
Query Match 3.8%; Score 100.5; DB 3; Length 518;
Best Local Similarity 19.9%; Pred. No. 0.11; Mismatches 139; Indels 177; Gaps 21;
Matches 94; Conservative 63; Mismatches 139; Indels 177; Gaps 21;
Db 93 PSVYFYRMDPIMWFLIGVPIVRHGPARKAFKEMNSASFIVYVVFVLLACYGVANN 152
130 PDFFKURIDDKNNIKISIGLI-----VVFPLIT-HSGFVG 168
Query 153 RIHTVMSLGPPEA-----VIAVFLVNLIP-----IMIPILW 187
Db 169 KU-----PESAFGLNYHAGLILIVAVIVFYIFGGYLAVSITDFQCVIMJAMM 219
Query 188 YEARKAKLFLWDDE-----VLYIQ 209
Db 220 VPIVALKLNGWDTFHDIAQMKPTNLDLFRGTVLGIUSLFSWGLGYFGQPHITVREMS 278
Query 210 ISGHSLPLKLRQKAV-YIATVPIISVLSWVIT-----HVTMSD-----LNINQVPP 255
Db 279 IKSHKHLKPKARRGIGSWMAGL-LGAIVGVLGIGSIFISERHKLEDPETLPIVMSQILP 336
Query 256 YCILDN--LTWIGAMWFLCEAMSTSIAHLLAERQKALKHIGPAMWADYRVLWLSK 313
Db 337 HPLVGGPLLAIAAAMSITSSQSLVTTSSLTEDFYKLTGSDKASSHQKEFVLRSLV 396

; ORGANISM: Drosophila melanogaster
US-09-491-577-98
Query Match 3.8%; Score 101; DB 4; Length 395;
Best Local Similarity 18.9%; Pred. No. 0.066;
Matches 56; Conservative 58; Mismatches 112; Indels 70; Gap 12;
Db 222 KAVVIALV-----PILSV-----ISVWTH-----VTMSDININQVPPY-- 256
97 KSFYIBILQSDDPPIINILVKEITRSLVLSIRNULMGCCCTCIGFVYIPFOSERVLPYGM 156
Query 257 -----CLUDNTLAMLGAWFLICEAMMSITAHLLAERQKALKH-IGP 297
Db 157 YLPTIDBEKYASPYEIFFVQIAIMAPMGCWYIPTVNWVTFALIMCRVLRKLS 216
Query 298 AAMVAVYR-----LW-----LRLSKLTDIGNALCYTFVFLMSLYFP--ITLSYGLMS 346
217 LEFLKEDQVRQBRBIWIKYQKLSFV-DSMNL-NTHLILVERFCFGAMLCVULSLL 274
347 QLSBGGFQKIDIGLITTAALWNGLLFLYFICDEAYA-----SVNRTNFOKLIM 394
Db 275 AQTIAQTVIVIAYMMIFANSVLYVANELYFQTRVWQSFKEFYKGILQSFIAAA 334
Query 395 VELWNWMSDACTEINM-FLRATEMNPSTINGGGFFDVNRTLFKGILTWTMVLV 449
Db 335 YESNWMDPVDVDTQTKLFLMRSQPLATLVGFTPMNLKMQSLNIAVSFPTL 390

RESULT 11 US-09-949-016-10076

Sequence 10076, Application US/09949016

PATENT NO. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastaSEQ for Windows Version 4.0

SEQ ID NO 10076

LENGTH: 1989

TYPE: PRT

ORGANISM: Human

US-09-949-016-10076

Query Match 3.8%; Score 99.5; DB 4; Length 1989;

Best Local Similarity 19.8%; Pred. No. 1.1;

Matches 102; Conservative 80; Mismatches 177; Indels 155; Gaps 25;

Qy 57 KYSIGLAERLDADYEAAPPLDRKKSADS-TASNNEPEKPSVYPRNIDPIN-WPARIIT---

Db 1361 KYRHCFPNTSERFPIEDVYKTCKEKMGEGNNNTBIRWKVKINPKFDNVGAGYIALLQAT 1420

Qy 111 -GULPIV---RHPKAKRKFENASAFISYVFFULLACIVGVYVANRRIHVSIG 162

Db 1421 FGKGMIDWAAVDSRKPDQPKYEDNIVYVYFIFR-----G 1460

Qy 163 PFEBAVAYIYLFLNTPIMTIPILWYEARKIAKLFNNDWDFEVLYQISGHSLPLKLR 222

Db 1461 SF-----FTLNUFIGVII-----DNEQQKKKEGGDQIMTEOK 1495

Qy 223 AVYTA-----IYLPILSVLSVWVTHVMSDLNINQVYPCILDNTMLG--- 267

Db 1496 KYNNAMKKGKSKKQPKPRPLNKQIYIVDFVQAFDIVIMMLCL--NMVTTMVEVD 1553

Qy 268 -----AWFLICEANSITIHALAERFOKAJGH---IGPAMWADYRVMWLSKL 314

Db 1554 TOSKOMENTLYWILWVFTFCVLUKME--ALRHYYFTG--WNIFDPVVVLSI-- 1606

Qy 315 TRDTGNAALCYTF--VFMSLVLFITLISYGLMSQSLSEGR-GIKD---IGLTITALWNI 367

Db 1607 --VGMFLADILKRYFVSPFLFVIRLARIGRLIKGAGTIRLFLAMMSLPALENT 1653

Qy 368 GILLEYTI-----CDEAHYASVNRTNFQ----- 396

Db 1664 GILLFLFLVMTTSIFGMSNFAVKGKHEGIDDMNFETFGNSMICLQFITSAGMDGCLLPI 1723

Qy 397 LANW--MNSQATEINMELRATEMPSTINGGEPDVNRTFKGLTTWVYLQRO 453

RESULT 12 US-09-538-092-564

Sequence 564, Application US/09538092

PATENT NO. 6753314

GENERAL INFORMATION:

APPLICANT: Mansfield, Traci A.

TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

FILE REFERENCE: 1566-542

CURRENT APPLICATION NUMBER: US/09/538,092

PRIOR APPLICATION NUMBER: 60/127,352

PRIOR FILING DATE: 1999-04-01

PRIOR APPLICATION NUMBER: 60/178,965

PRIOR FILING DATE: 2000-02-01

NUMBER OF SEQ ID NOS: 1187

SOFTWARE: CurateSeqFormatter Version 0.9

SEQ ID NO 564

LENGTH: 1900

TYPE: RT

ORGANISM: *Saccharomyces cerevisiae*

FEATURE: misc_feature

NAME/KEY: misc_feature

LOCATION: (0)..(0)

OTHER INFORMATION: Polypeptide Accession Number YLR305C

US-09-538-092-564

Query Match 3.7%; Score 99; DB 4; Length 1900;

Best Local Similarity 19.2%; Pred. No. 1.2;

Matches 90; Conservative 66; Mismatches 154; Indels 158; Gaps 20;

Qy 70 YEAPPDRKKSSTASNNPERKPSVFTRNIDPINFPLRIGVPIVHGP----- 122

Db 727 YNSPPL-----ASERPPANKEL-----LEMTTILRKGSSENNIQOK 764

Qy 123 -----AKKEMNSAS--PIYSVFPVFLACIVGVYVANRRIH--VRSLSGP----- 166

Db 765 QQTBYFNTNIVQRTTSSSKIMLAATLETIRCGACDKSNTLYPSDPSLISGSTEK 824

Qy 167 AVIAYLFLNTPIMIPIIYLWYEARKIAKLFNNDWDFEVLYQISGHSLPLKLR-- 220

Db 825 CIA-----VLSVSMI-----RKYARIQKONDADFNNSKMQAQQLNLICLISHRP 870

Qy 221 --QKAVYIAYIYLFLNTPIMIPIIYLWYEARKIAKLFNNDWDFEVLYQISGHSLPLKLR----- 278

Db 871 TLQDAKFAHCEIFIRSPSSLCHHL-----VYTLDMUTLAFDS-----ILD 913

Qy 279 ITAHLLAERFOKAJGHGPMW--ADYRV-----WLR--LSKUTDTGNA 322

Db 914 SEAHKPRPRFKGKSTTILVPPSSRATLSRLHKSAKENRILNRSNODT--- 969

Qy 323 CYTFVMSLXLFITLISYGLMSQSLSEGR--FGIKDGTITALWNIIGLFLYICBHYA 380

Db 970 -----KILQSYISDGLYSRSLNSVFGVSPAMDGLPALKELSLRITYGP----- 1019

Qy 381 SUNVRTNFQKLMVTELNNMSDAQTEINPLRATEMPSTINGCGFPDVNRTFKGLUT 440

Db 1020 -----KENTIS--GFLSLHMSRSKLF----- 1040

Qy 441 TMVTLVLUQROQSIPTD--KGSEGANNITVUDPFDMSLNDMSL 486

Db 1041 TATSSPEDIKQICISTONKRNLTGKKTQD-VTDFDMATALL 1087

RESULT 13 US-09-489-039A-8212

Sequence 8212, Application US/09489039A
; PATENT NO. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEARIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709_200401
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: -US 60/117, 747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO: 8212
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-8212

Query Match 3.7%; Score 98.5; DB 4; Length 656;
Best Local Similarity 21.2%; Pred. No. 0.26; Mismatches 26; Indels 165; Gaps 28;
Matches 118; Conservative 67; Mismatches 206; Indels 165; Gaps 28;

Qy 32 KGDADVFLNAKP----LNSANQAQAYLGVRKSIGLAERLDADY--EAPPDRKKSSDST 84
Db 75 KADAESEKIKANPQGVIMVYESGQFQVQVIGNHFDWMLAVRQAGLSDSEPVAAERKA-- 131
Qy 85 ASNNPBPKPSPVYRNIDPINWFLRITIGVLPIV-----RHGPRAK 125
Db 132 -----KGSVLSQOLIDISGIFTPIGVMATGMLKGLLALLAVTCGWLTPEQGYKIKW 184

Qy 126 ENSASPIYSVTFEVLLACYGVYVANNR----HIVRSLSGPFEAEAVIAYFLVNLIP-- 179
Db 185 ASDALFF----FFPL----FLGYTAGKFGGGNPFMSVIGGALTHPLMIAFEASOPGA 237

Qy 180 --IMIPIWM--YEARKIAKLNDQWDDPFLVYQISGHSPLKUR--QKAVVIAVL 230
Db 238 AVEHFLGIPVTFINKESSVILASW--VSCWLERKSNAIPLPSMKNFSPACIUVV 295

Qy 231 PI-LSVLSVWVTHVMSDLNINQVQVYPCILDNLT-AMIGAWWF-----LICE 275
Db 296 PLTFLYIGPVWTLHSLHANGYQFI-YAFAWLWLAGVIGAMQVCVIFGLHWWGLYPLMIN 354

Qy 276 AMSTIAHTLAAFRQOKALKHIGPAAWADYRVTWLRSLSKLTDTGNALCYTFVFMILYLF 335
Db 355 NWTVLGH----DSMPLITFLAVIAQVAVGVLGIFLA--TRDARORVLAGSAF-SAGLFG 405

Qy 336 LITLSTYGLMSQLSBCP----GKDIGUTTAL----WNIGULFYICDEAHAYASVNU 384
Db 406 ITEPAIYGLTFLRRLPPIFGCVAAGA--IGGAATAFNSVASYFGL----- 448

Qy 385 RTNFQKLLMVELNWNMSDAQTEINMPLRATENMPSTINCQGFFDVNRTLFKGLLTMT 444
Db 449 -----PNTFPQAMIPP----GG--IDASWVGLGIGTGVA 477

Qy 445 Y-LVVLVQFOFISIPTDKGDSIGANNITVUDFPMDSIDNDMS-----IM 486
Db 478 FVLAQCVLTFFAGLPR--GSAPAGAVTVAPASANDILAPMSGVIALEQVFDSTFASGLL 534

Qy 487 GASTLSTTVCGLTLP 502

Db 535 GRGVATIPAVGQVIA 550

RESULT 14
US-08-671-525B-6
; Sequence 6, Application US/08671-525B
; PATENT NO. 5703220
; GENERAL INFORMATION:
; APPLICANT: Yamada, Tadataka
; APPLICANT: Gantz, Ira
; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

RESULT 15
US-08-672-109B-6
; Sequence 6, Application US/08672109B
; PATENT NO. 5710265
; GENERAL INFORMATION:
; APPLICANT: Yamada, Tadataka
; APPLICANT: Gantz, Ira
; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: US
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671, 525B
FILING DATE: June 27, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Dean F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-000853DVB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 360 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
; US-08-671-525B-6

Query Match 3.7%; Score 98; DB 1; Length 360;
Best Local Similarity 21.3%; Pred. No. 0.12; Mismatches 85; Conservative 64; Mismatches 172; Indels 78; Gaps 17;
Matches 85; Conservative 64; Mismatches 172; Indels 78; Gaps 17;

Qy 32 KGDADV----LNAPEKLNNSANQAQAYLGVRKSIGLAERLDADYEAPPD 76
Db 9 EGDGFVPUVSSSPFLRTLPLQPGSALTMAMASCCUPSPVQPTLPNGSEHLQAPFFS---- 64

Qy 77 RKKSSDSTASNPNPEKPSVFRNIDPINWFLRITIGVLPIVHGPRAK 136
Db 65 --NOQSSAFCQVFKEPEL-SLGIVSLLNENLIVLAVTRANGNLHSPMF----FLCSL 117

Qy 137 VPFVLLACYGVYVANNRHTAVRSLSGPFEAEAVIAYFLVNLIPMIPITIYEARKIAKL 196
Db 118 AVADMLVSVSALETTMIAVHSVDTFEDOFOH--MDNIFDSMCISL-----VASI 169

Qy 197 FN----DWDPEFLVYQISGHSPLKURQKAVVIAVFLPLSLSVWVTHVMSDLNINQ 252
Db 170 CNTLATAVDRYTFYALRHISI-MTVRKALTLIVAIWVCCGCVGUVF--IVYSE---SK 223

Qy 253 WPYCILDLNLTAMLGAWWWFLICBAMSITAHIAERQOKALKHIGPAAWADYRVLRLS 312
Db 224 MVIUCVLTIMPFAMM----LMLGTVLWMLFARHVKAFAALPPADGVWAPQHSCMKGA 278

Qy 313 KLTDRDGNALCYTFV-----MSLYVLPFTITISIYGLMSQSEGFGKDIDGTTIAWNNG 368
Db 279 V----TITILGVIFRCWAPPFLHLVLIITCP-----TNPYCICYTAHFTY 321

Qy 369 LLYFTC----DEAHYA--SIVNRTFQKGLMVELNWNM 401
Db 322 LVLIMCNSVIDPLTVAFRSLBRNTPRE--ILCGCNGMN 358

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
 STREET: P.O. Box 828
 CITY: Bloomfield Hills
 STATE: MI
 COUNTRY: US
 ZIP: 48303

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08672.109B
 FILING DATE: June 27, 1996
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Deann F.
 REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-000853DVC

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (810) 641-1600
 TELEFAX: (810) 641-0270

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
 LENGTH: 360 amino acids

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-672-109B-6

Query Match 3.7%; Score 98; DB 1; Length 360;
 Best Local Similarity 21.3%; Pred. No. 0.12; Matches 85; Conservative 64; Mismatches 172; Indels 78; Gaps 17.

Matches 85; Conservative 64; Mismatches 172; Indels 78; Gaps 17.

Qy 32 KGDAVE-----|NAPKLNAAAQAYLYGVKYSIGLAERIDADYAPPLD 76
 Db 9 EGDIVFPVSSSFLLETLERQQLSALLTAMASCLLPSVQPTLPGSEH-QAPFFS--- 64

Qy 77 RKKSSTDSSTASNNEPKPSVYRNPIDPINDPWNRLRIGSYLPTVHGRPARKEFKNSAARIYV 136
 Db 65 --NQSSSAFCEQVFKPEIFL-SLGIVSLLNLYLTLAVWRNGNHSMPY---FLCL 117

Qy 137 VFFVLLACTGYVANRRIHVRSLSGPFEAAVIANFLVNLPIIMIPIUVEARKIAKL 196
 Db 118 AVADMLVSVSNALETTIMIAVHSVDLTFEPQFOI--MDNIFDSMCISL-----VASI 169

Qy 197 FN---DWDPPFEVLYTOISGHSLPPLKLRKAVYIAVLPILSVLSSVITHVMSDLINQ 252
 Db 170 CNLLIAVDRYVTFYALRHSI-MTVRKALTLIAIWCVCGVCGVVF--IVSE--SK 223

Qy 253 WPPYCLDNLNTAMIGGAAWFLICEAMNSITAHLAERQKALKHIGPAMVADYVRLRLS 312
 Db 224 MVIYCLITNEFAMM---LIMGTYLXVHMFELFARLAVKRTIALPAGDVQAOQHSCMKGA 278

Qy 313 KLTDRDGNALCYTFVF---MSLYVFFITISIYGLMSQLSSEGFGJKDIDGITLTAWNG 368
 Db 279 V---TITILGVFVFCWAPPLFLHLVLTICP-----TNPYCICYTAHFTY 321

Qy 369 LLFVYC---DEAHYA---SNTVRTFQKQKLMVLELMNN 401
 Db 322 LVLINCNNSVTDPLVAFRSKBLRNTRE--IICGCMGN 358

Search completed: May 23, 2005, 19:26:45
Job time : 46 secs

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

May 23, 2005, 19:19:25 ; Search time 135 Seconds
 (without alignments)

1268.651 Million cell updates/sec

Title: US-10-081-816-12
 Perfect score: 2645
 Sequence: 1 MRPSGEKVKVKGHGCGNSGHS...TTVGGTTLPPPIKLUKGKKG 512

Scoring table: BLOSUM62
 GapP 10.0 , Gapext 0.5

Searched: 1434725 seqs, 33450795 residues

Total number of hits satisfying chosen parameters:

1434725

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0*
 Maximum Match 100*
 Listing first 45 summaries

Database : Published Applications_AA.*
 1: /cgn2_6/ptodata/2/pubpea/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpea/PCT_NEW_PUB_pep:*

3: /cgn2_6/ptodata/2/pubpea/US06_NEW_PUB_pep:*

4: /cgn2_6/ptodata/2/pubpea/US05_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpea/US07_NEW_PUB_pep:*

6: /cgn2_6/ptodata/2/pubpea/PCUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpea/US08_PUBCOMB.pep:*

8: /cgn2_6/ptodata/2/pubpea/US09_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpea/US09B_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpea/US09C_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpea/US09D_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpea/US09A_PUBCOMB.pep:*

13: /cgn2_6/ptodata/2/pubpea/US09B_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpea/US10C_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpea/US10D_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpea/US10C_PUBCOMB.pep:*

17: /cgn2_6/ptodata/2/pubpea/US10D_PUBCOMB.pep:*

18: /cgn2_6/ptodata/2/pubpea/US11C_PUBCOMB.pep:*

19: /cgn2_6/ptodata/2/pubpea/US11D_PUBCOMB.pep:*

20: /cgn2_6/ptodata/2/pubpea/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

14 120 4.5 450 14 US-10-081-816-57 Sequence 57, Appl
 15 119 4.5 477 15 US-10-47-328-70 Sequence 70, Appl
 16 118.5 4.5 2008 9 US-09-736-969A-2 Sequence 2, Appl
 17 118.5 4.5 2008 9 US-09-736-969A-91 Sequence 91, Appl
 18 118.5 4.5 2008 9 US-09-737-246-99 Sequence 99, Appl
 19 118.5 4.5 2008 9 US-09-736-960-88 Sequence 88, Appl
 20 118.5 4.5 2008 9 US-09-736-968A-105 Sequence 105, Appl
 21 118.5 4.5 2008 10 US-09-918-24A-29 Sequence 29, Appl
 22 115 4.3 367 14 US-10-081-816-13 Sequence 13, Appl
 23 112.5 4.3 365 14 US-10-081-816-56 Sequence 56, Appl
 24 112.5 4.3 408 15 US-10-447-328-50 Sequence 50, Appl
 25 110.5 4.2 408 14 US-10-081-816-37 Sequence 37, Appl
 26 110.5 4.2 2073 10 US-09-918-24A-10 Sequence 10, Appl
 27 110 4.2 498 14 US-10-081-816-15 Sequence 15, Appl
 28 108 4.1 381 15 US-10-447-328-34 Sequence 31, Appl
 29 107.5 4.1 346 15 US-10-282-122A-77576 Sequence 34, Appl
 30 107.5 4.1 404 14 US-10-081-816-52 Sequence 77576, A
 31 107 4.0 390 14 US-10-081-816-39 Sequence 52, Appl
 32 106.5 4.0 390 15 US-10-447-328-58 Sequence 39, Appl
 33 106.5 4.0 702 15 US-10-474-776-384 Sequence 58, Appl
 34 106.5 4.0 702 17 US-10-472-928-4070 Sequence 384, App
 35 106.5 4.0 816 10 US-09-232-089-5 Sequence 4070, Ap
 36 106.5 4.0 816 10 US-10-340-792-5 Sequence 5, Appl
 37 106.5 4.0 816 14 US-10-081-816-52 Sequence 5, Appl
 38 104 3.9 363 15 US-10-601-309-74 Sequence 74, Appl
 39 103.5 3.9 960 14 US-10-234-390-4 Sequence 4, Appl
 40 103 3.9 287 14 US-10-081-816-22 Sequence 22, Appl
 41 103 3.9 1011 15 US-10-300-797-3324 Sequence 324, Ap
 42 102.5 3.9 373 14 US-10-081-816-4 Sequence 4, Appl
 43 102.5 3.9 373 15 US-10-447-328-8 Sequence 8, Appl
 44 102 3.9 320 10 US-09-510-332-160 Sequence 150, App
 45 102 3.9 320 17 US-10-962-365-160 Sequence 160, App

ALIGNMENTS

RESULT 1
 US-10-081-816-12

; Sequence 12, Application US/10081816

; Publication No. US20030045472A1

; GENERAL INFORMATION:

; APPLICANT: Axel, Richard

; SCOTT, Kristin

; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recepto

; TITLE OF INVENTION: Uses Thereof

; FILE REFERENCE: 0575/64019-AJ/PW/ADM

; CURRENT APPLICATION NUMBER: US10/081,816

; CURRENT FILING DATE: 2002-02-22

; PRIOR APPLICATION NUMBER: 60/271,319

; NUMBER OF SEQ ID NOS: 116

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO: 12

; LENGTH: 512

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; US-10-081-816-12

; Query Match

; Best Local Similarity

; 100.0%; Score 2645; DB 14; Length 512;

; Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Sequence 12, Appl

; Sequence 24, Appl

; Sequence 10, Appl

; Sequence 1, Appl

; Sequence 67, Appl

; Sequence 80, Appl

; Sequence 29, Appl

; Sequence 21, Appl

; Sequence 30, Appl

; Sequence 32, Appl

; Sequence 11, Appl

; Sequence 30, Appl

; Sequence 1, Appl

; Sequence 2, Appl

; Sequence 1, Appl

; GENERAL INFORMATION:
; APPLICANT: Axel, Richard
; APPLICANT: Scott, Kristin
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor
; FILE REFERENCE: 05/5/64019-AJJPW/ADM
; CURRENT APPLICATION NUMBER: US/10/081,816
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO: 67
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-081-816-67

Query Match 6.4%; Score 169; DB 14; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.3e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 33; Conservative 0; MisMatches 0; InDelS 0; GapS 0;

Qy 420 STINCGGFDPDVNLKFGKLTMTVYLVQF 452
Db 1 STINCGGFDPDVNLKFGKLTMTVYLVQF 33

RESULT 5

US-10-447-328-80
Sequence 80, Application US/10447328
Publication No. US2004003419A1

; GENERAL INFORMATION:
; APPLICANT: Carlson, John R.
; APPLICANT: Clyne, Peter J.
; APPLICANT: Marr, Coral G.
; APPLICANT: Yale University
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor
; FILE REFERENCE: 05/5/64019-AJJPW/ADM
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 29
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-081-816-29

Query Match 5.1%; Score 135.5; DB 14; Length 369;
Best Local Similarity 20.5%; Pred. No. 0.0021; Mismatches 65; Indels 147; Gaps 20;
Matches 87; Conservative 65; MisMatches 126; InDelS 147; GapS 20;

Qy 106 FLRIGVUPVIRGPARAKFEMNSASFYSVWVFVLLACYVGVANNR-IHIVS-165
Db 11 YFALLGLVP-WSSBSCAQSKPVK---VYSAILIL---NAVFGISIYFP-Q 54

Qy 166 EAVIAYLFLVNL---PMIITLWYARKIAKLFNWDDF---FVLYY---QI 210
Db 55 SAEFLSLMVNVIVFVAVCVCVVI---ILQVMMVDDYFRCRMEKYGLRLQCR 108

Qy 211 SGHSILPLKURQKAVYTAI-----VLPPTL-----SVLSVVIT-----241
Db 109 KIHVGRLKNSQYAKILALGIGFVTVL-----SIVVALSGSIIYFWSLISLILRQFVYTL 168

Qy 242 -----RHTMSDLNTQW-----PCILDNLTTAMGAWLTCIEMS-ITPAHLER 287
Db 169 NVELLGHHRVSLGIGRLQNTLCHLMGANTTDGNGANRCLSPLALQKSHMQLHLYLTH 228

Qy 288 PQKALKHIGPAAWYDVRVWLRLSKLRTDGTGALCYTFVPMISLYRPFITISYGLMSQ 347
Db 229 FNDLP-----GMSILGTV---VLFSSTDSTVNYWTQCV 258

Qy 348 LSEGGPKDQIGLT---ITPAWNGLLFLVICDE-----AHYASVNRVN 387
Db 259 LVEVTEKYLATTSVFTSPFPN1-LVFGRCGRCQRSVLSYRLNLSCHASIGRTS 317

Qy 388 FOKKLMVNELNWMSDADTEINMLRATENPSPN1NC3CFDVFNTLFPKGLLTMVTV 447
Db 318 Y-KDLM-----EPILQEVQVLAIAEFGMSTONSLMSILAKVTLI 361

Qy 448 VLUQF 452

Query Match 6.0%; Score 160; DB 15; Length 433;
Best Local Similarity 20.9%; Pred. No. 1e-06; Mismatches 139; Indels 134; Gaps 18;
Matches 91; Conservative 72; MisMatches 139; InDelS 134; GapS 18;

Qy 102 PIMFLRIGVLPVIRGPARAKFEMNSASFYSVWVFVLLACYVGVANNR-IHIVS-159
Db 27 PIOPFTRTL-----HKRRGIGTIVLGYACVYLISIISLIVYECTVANTVQKQJHKFHAE 79

Qy 160 ---LSPPF-BEAIVAYLFLVNLPIIMIPLWYARKIAKLFNWDDEV-LYQISG-212
Db 80 DSSKMGKFAERSWVWAMFVNQNLILL---NFRRLARIYDIADELDIINNSACP 133

Qy 213 ---HSLPLKURQKAVYTAI-----VLPPTL-----SVLSVVIT-----241
Db 134 VGQRHWWRFRFLAISVGLWVLLVGLTPFTLVALGVPLHWTNKULTEILIMQL-KC 192

Qy 254 VPYCILDNLTAMLGAWMFLICEAMSITAHLLAERFQKAKHIGPAMVADYVRLWLRISK 313

Db 193 TEYCV-----FVLLVYELILGRHIL-QQISVLE-----GN 223

RESULT 7
 US-10-081-816-17
 Sequence 17, Application US/10081816
 Publication No. US20030045472A1
 GENERAL INFORMATION:
 APPLICANT: Axel, Richard
 TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And olfactory Receptor
 TITLE OF INVENTION: Use Of Thereof
 FILE REFERENCE: 0575/64019-A;JPW/ADM
 CURRENT APPLICATION NUMBER: US/10/081,816
 CURRENT FILING DATE: 2002-02-22
 PRIORITY FILING DATE: 2001-02-23
 PRIORITY FILING DATE: 2001-02-23
 NUMBER OF SEQ ID NOS: 116
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 17
 LENGTH: 736
 TYPE: PRT
 ORGANISM: *Drosophila melanogaster*
 US-10-081-816-17
 Query Match 5.0%; Score 132.5; DB 14; Length 736;
 Best Local Similarity 21.0%; Pred. No. 0.0011; Mismatches 80; Conservat 80; Mismatches 135; Indels 89; Gaps 17;
 Matches 77; Conservative 66; Mismatches 134; Indels 89; Gaps 16;
 QY 102 PTFNWLRLIGVPIRIVKQPARAKEEMASASFYSVVFLVLLACVYGVYANNR-IHIVRS- 159
 QY 427 PIQPFTRPL-----HKRRRGIVTLGACYCILISLIVYCYANTVALQDKIKFHR 479
 QY 160 ---LSSGFEEAVIAFLFLVNLIPIMIPLWYEARKEIAKLFNDWDDPEV-LYQISG- 212
 QY 480 DSSKGWMTQKULVWAMVWNLIL-----NFRLLAVIDIADLEIDANNASSFV 533
 QY 213 --HSLPLKLQR-QKAVVIAVL-----PILSVLWVVTAVTMSDNLNQVY 254
 QY 534 GQRHWWPRFRFLALISVGLWIVLVLGLTRPLFTVALGPYIHWTKVLTIELIMLQL-KCT 592
 QY 255 PYCILDLTATMAGAWMFLICEAMNSITAHLLAKRQKALKHIGAAMVADYVWLRASKL 314
 QY 593 EYCV-----FVLLYLIELTRGRHLI-QQISVELE-----GNQ 623
 QY 315 TRDTGNAICYTFFVMSLFLFITSIYGLMSQSEGCGKIDGTTALMWNLGLFFIC 374
 QY 624 SDSVPSLC-----VALKRNLQLAGR,WGLVNEVSY-----TSLTFLYNTLITQIV 675
 QY 375 DEAHAYASVNR--TNFOKLULMELNNMSDAQTEINMFLRATEMPSTINCQGFDVNR 432
 QY 676 NWALKSYSPNNECCQVTDYLILKQREYSLQMEHLKL-----PTCGGLFDNL 726
 QY 433 TLFKGL 438
 QY 727 KFPGGV 732
 RESULT 8
 US-10-081-816-30
 Sequence 30, Application US/10081816
 Publication No. US20030045472A1
 GENERAL INFORMATION:
 APPLICANT: Axel, Richard
 TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And olfactory Receptor
 TITLE OF INVENTION: Use Of Thereof
 FILE REFERENCE: 0575/64019-A;JPW/ADM
 CURRENT APPLICATION NUMBER: US/10/081,816
 CURRENT FILING DATE: 2002-02-22
 PRIORITY FILING DATE: 2001-02-23
 PRIORITY FILING DATE: 2001-02-23
 NUMBER OF SEQ ID NOS: 94
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 32
 LENGTH: 372
 TYPE: PRT
 ORGANISM: *Drosophila melanogaster*
 US-10-081-816-30
 Query Match 4.9%; Score 130.5; DB 15; Length 372;
 Best Local Similarity 21.7%; Pred. No. 0.00067; Mismatches 80; Conservat 80; Mismatches 135; Indels 89; Gaps 17;
 Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17;
 QY 129 SASFTYSVVFVLLACVYGVYANNRHHI-VRSLSG-PFEEAVI-----AVFL-- 174
 QY 39 STAQIQRIVVGVFMALGAGLAEISLYMETSQQTNTFONAVILUTSVTQLLNLWRSQ 98
 QY 175 ---VNILPIM--TIPILWYARKTAKLFENDDDPEVLYQQISQGHSPLKROKAVYAI 228
 QY 99 QRSQVNLLQRLSQVELLQFPIAVQ-----FRMLV-----RI 132
 QY 229 VPLTISLVSLSVUTHTVMSDNLNQV-----VPCYCLDNLNTAMLGAWMFLICEAMSI 279
 QY 133 WLYCLLIGVAMWTHRGVNLWTLQISRSVTLIGFVRCVLN-----FOPTCYTGMV 185
 Db 280 TAHLLERFQALKI-----GPAHADYFVWMLRSKTDTOMALCUTPVF- 328
 Db 186 ILLKLUQDQVQKQLBLVSTTISAGWAGCRLTDEILQGORELTAVYGVWPLFLFIQ 245
 Db 329 --MSLYLFFTITLSTYGLMSQLSFGFGKIDGTTALMWNLGLFYICDEAHYASVNR 386
 Db 245 WMQCLLFTIS-----EGHSN-DMIFCMLAIPMFLYFLWVNDHQA 294
 QY 387 NFQKLUMLTENNNMSDAQTEINMFLRATEMPSTINCQGFDVNR 445
 Db 295 NTKAKML-TKPVRTGIGLDRMIEKFLKNRQKPLTAVGFALDKSTLKF-LFTAIFTY 352
 QY 446 LWLJLQFQ 453
 Db 353 MVLVQFK 360
 RESULT 9
 US-10-0447-32B-32
 Sequence 32, Application US/10044732B
 Publication No. US20040003419A1
 GENERAL INFORMATION:
 APPLICANT: Carlson, John R.
 APPLICANT: Clyne, Peter J.
 APPLICANT: Warr, Coral G.
 APPLICANT: Yale University
 TITLE OF INVENTION: US20040003419A1 Taste Receptors in *Drosophila*
 FILE REFERENCE: 44574-5072
 CURRENT APPLICATION NUMBER: US/10/0447,328
 CURRENT FILING DATE: 2003-05-29
 PRIORITY FILING DATE: 2000-06-14
 PRIORITY APPLICATION NUMBER: US/09/593,519
 PRIORITY FILING DATE: 2000-06-14
 PRIORITY APPLICATION NUMBER: US 60/138,668
 PRIORITY FILING DATE: 1999-06-14
 PRIORITY APPLICATION NUMBER: US 60/181,704
 PRIORITY FILING DATE: 2000-02-10
 NUMBER OF SEQ ID NOS: 94
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 32
 LENGTH: 372
 TYPE: PRT
 ORGANISM: *Drosophila melanogaster*
 US-10-0447-32B-32
 Query Match 4.9%; Score 130.5; DB 15; Length 372;
 Best Local Similarity 21.7%; Pred. No. 0.00067; Mismatches 80; Conservat 80; Mismatches 135; Indels 89; Gaps 17;
 Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17;
 QY 129 SASFTYSVVFVLLACVYGVYANNRHHI-VRSLSG-PFEEAVI-----AVFL-- 174

RESULT 10
 US-10-081-816-11
 ; Sequence 11, Application US/10081816
 ; Publication No. US2003045472A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Axel, Richard
 ; APPLICANT: Scott, Kristin
 ; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor
 ; TIME OF INVENTION: Uses Thereof
 ; FILE REFERENCE: 0575/64019-AJPW/ADM
 ; CURRENT APPLICATION NUMBER: US/10/081,816
 ; PRIORITY APPLICATION NUMBER: 60/271,319
 ; PRIORITY FILING DATE: 2001-02-23
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 11
 ; LENGTH: 436
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; US-10-081-816-11

Query Match 4.9%; Score 129; DB 14; Length 436;
 Best Local Similarity 21.2%; Pred. No. 0.0012; Mismatches 172; Indels 108; Gaps 22; Matches 94; Conservative 70; Mismatches 172; Indels 108; Gaps 22;

QY 85 ASNNEPEFKP-SVFRYRNDPTNFMFLRIGVLP--IVRHGPRAKPEMNSASFVSVVFFV 140
 QY 30 AQQGLEFEGQDFTGAIKPVLCVAOPFGIMPLNSWRSRQDVKRSIGAVTGLELL 89
 QY 141 L--LACYGVGVANNRHHIVSLSRPFEEAVIAYLFLVNLIPIMIPILYME--ARKIA 194
 QY 90 LGGMKTLVGLG--AN---ILFTEGLNANQVNLGVFLI----VGVNNNLNFGEARSWS 136
 Db 195 KLFNUWDDDEVLY--YQISGHSLPKRQKAVVIAVLPILSY----- 235
 Db 137 HIMLPWSSVDTIMLFPYKRGKRS---LRSKVNTLALSVVLAVGDHMLYYASGCSYS 192
 QY 236 LSVVTHVTMSDL-----NINQVVFYCLDNLNTAM-LGAW-----WFLICE 275
 Db 193 MHIQLOCHTWSRITFGYLVLEKFSLIMFMP---NITSMCYGWLNGAFTFLWNFMFI 248
 QY 276 AMSITAHLLARFORKALKHIG-----PAMVADYRVMLSLSKLTDRTGNCALCYTFV 328
 Db 249 FIVMISIGLAQRFQFOFAARVGALEGHRVPEALWDIRDHIRLCELA-----SIVEAS 301

RESULT 11
 US-10-447-338-30
 ; Sequence 30, Application US/10447328
 ; Publication No. US2004003419A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Carlson, John R.
 ; APPLICANT: Clyne, Peter J.
 ; APPLICANT: Warr, Coral G.
 ; APPLICANT: Yale University
 ; TITLE OF INVENTION: No. US2004003419A1el Taste Receptors in Drosophila
 ; FILE REFERENCE: 4474-5072
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIORITY FILING DATE: 2000-06-14
 ; PRIORITY APPLICATION NUMBER: US/09/593,519
 ; PRIORITY FILING DATE: 2000-06-14
 ; PRIORITY APPLICATION NUMBER: US 60/138,668
 ; PRIORITY FILING DATE: 1999-06-14
 ; PRIORITY APPLICATION NUMBER: US 60/181,704
 ; PRIORITY FILING DATE: 2000-02-10
 ; NUMBER OF SEQ ID NOS: 94
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 30
 ; LENGTH: 348
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; US-10-447-338-30

Query Match 4.8%; Score 127; DB 15; Length 348;
 Best Local Similarity 19.8%; Pred. No. 0.0013; Mismatches 129; Indels 138; Gaps 17; Matches 81; Conservative 62; Mismatches 129; Indels 138; Gaps 17;

QY 106 PLRIGVLPVIRHGPRAKPEMNSASFVSVVFFVLCY----- 145
 QY 11 YFALIGLVP-WESCAQOSFVK---YSAILIILVHFGISIYFPOSAELFLSLMG 65
 Db 11 YFALIGLVP-WESCAQOSFVK---YSAILIILVHFGISIYFPOSAELFLSLMG 65

QY 146 -VGTV-----ANRHTIVSLSLGPPEEAV-IAYLFLVNLIPIMIPILYE 189
 Db 66 GIGFVTRIAGTYGLRLOCBLKHVGKRWQSYAKIALGIGLVLTLPISVAL--- 121

QY 190 ARKIAKLENDWDDDEVLY--YQISGHSLPKRQKAVVIAVLPILSY----- 249
 Db 122 -----SGSILYFWSSLISI-LIRMOFLVVLINVELLG---HRVSLGJIR 162
 QY 250 INQV-----PYCILDNLTAMLGAWWFLICEAMS-ITAHLLARFORKALKHIGPAAMVA 302
 Db 163 LQNLTECHLGMGCTLDGMANRICKSLEFLALKOSHMLYHFLPHNLF----- 212
 QY 303 DYRVVLLRSLKLTDTGNCALCYTFVMSLXLFITLSTYGLMSQLSBGPGIKIGLT-- 360
 Db 213 -----GWSLIGTV---VLFSTVWVWYDQVWVYEVKLYATPS 252

QY 361 --ITALWNLGFLYFICDB-----AHYASVNRNTFQKGLMLVBNWMS 402
 Db 253 VFVPSFFN1-LVFCRGECFQROSVLISLGSYLRNLNSCHPSIGRTSY-KDLM----- 302
 QY 403 DAQTEINFLRATENPSTINGCERFVDFVNTLFLKGLLTMTWVLUQF 452
 Db 303 -----ERFLQVQVNLAINAEGMSTUNSLSMSILAKKTVVILVNG 345

RESULT 12
 US-10-081-816-1
 ; Sequence 1, Application US/10081816
 ; Publication No. US20030045472A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Axel, Richard
 ; APPLICANT: Scott, Kristin
 ; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor
 ; FILE REFERENCE: 0575/64019-A/JP/ADM
 ; CURRENT APPLICATION NUMBER: US/10/081,816
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR FILING DATE: 2001-02-23
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: Patentin version 3.1
 SEQ ID NO 1
 LENGTH: 410
 TYPE: PRT
 ORGANISM: *Drosophila melanogaster*
 US-10-081-816-1

Query Match 4.6%; Score 122; DB 14; Length 410;
 Best Local Similarity 22.2%; Pred. No. 0.0053; Mismatches 104; Indels 88; Gaps 18;
 Matches 70; Conservative 53; MisMatches 104; Indels 88; Gaps 18;

Qy 185 ITW-YEARKI---AKLFNDWDPEVLYQTSGHSLPL--KUROKAVYIAIVLPILSVL 236
 Db 153 ILWGVAVSQQLLTGAKLISRGDRFPI--YWIS-YIPLVCGIYRQINATOL-VRQL 208
 Qy 237 SVVITHVMSDIAINQVPCYCLDNLTAMLGAWMFLICEAMSITAHLAERFQALKHIG 296
 Db 149 ILWGVAVSQQLLTGAKLISRGDRFPI--YWIS-YIPLVCGIYRQINATOL-VRQL 204
 Qy 237 SVVITHVMSDIAINQVPCYCLDNLTAMLGAWMFLICEAMSITAHLAERFQALKHIG 296
 Db 149 ILWGVAVSQQLLTGAKLISRGDRFPI--YWIS-YIPLVCGIYRQINATOL-VRQL 204
 Qy 297 PAAWADYRVVWLRSLKLTTRDGNALCYTFVMSLVL----FFITLSTYGLMSQSEG 351
 Db 243 AVRLV-YQRMWALVALLNR----CYG--LSMLMQVNDPLAITSNCYWMFLNFRQS 291
 Qy 352 ---FGIKDIGITALLW----NIGLFLYICDEA-----HYASVNRNTNPK 390
 Db 292 AASPFDI--LQIVASGWSAHLGNVNLISLICDRTAQACASRLCHQVSVDLNESHN 349
 Qy 391 KULMVELNWMNSDAQTEINMPRATEKNPSTINCAGGFDVNTLFKGLLITWVTVLVL 450
 Db 350 ALI-----TQFSLQLLHORLH--PSAAGFFNVDCTLYTIVGATTWLL 394
 Qy 451 QFOQISIPTDKGSEG 465
 Db 395 QPHMSESTIGSDSNG 409

RESULT 13
 US-10-447-328-2
 ; Sequence 2, Application US/10447328
 ; Publication No. US2004003419A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Carlson, John R.
 ; APPLICANT: Clyne, Peter J.
 ; APPLICANT: Warr, Coral G.
 ; APPLICANT: Yale University
 ; TITLE OF INVENTION: No. US2004003419A1el Taste Receptors in *Drosophila*
 ; FILE REFERENCE: 44574-5072
 ; CURRENT APPLICATION NUMBER: US/10/447,328
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: US/09/593,519
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/138,668
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 60/181,704
 ; PRIOR FILING DATE: 2000-02-10
 ; NUMBER OF SEQ ID NOS: 94
 ; SOFTWARE: Patentin version 3.1
 SEQ ID NO 57
 LENGTH: 450
 TYPE: PRT
 ORGANISM: *Drosophila melanogaster*
 US-10-081-816-57

Query Match 4.5%; Score 120; DB 14; Length 450;
 Best Local Similarity 19.0%; Pred. No. 0.0096; Mismatches 149; Indels 162; Gaps 26;
 Matches 91; Conservative 76; MisMatches 149; Indels 162; Gaps 26;

Qy 71 EAPPDRKKSSTASNN-----PEP-KPSVYRNIDPINFRLIRIGVPIV--RHG 119
 Db 28 ETPP--PKFVEDSNLBNFLASEKPLNNTNLDFHRAVFPFMIAQCVAIMPLVGIRESN 85
 Qy 120 PARAKEEMNSASPIYVVFV----LLACY----VGVYANRRIHIVRSLSGPPEAEVI 169
 Db 86 PRRVFAYKSIPMFVUFLIMATSILFLSMETHLKIGITKVNFGVLF----FGCVL 140
 Qy 170 AYFLVNLIPIMIPIWYEARKIAKLFNDWDPEVLY---YQISGHSLP----- 216
 Db 141 AYVVFIRL-----AKKWPWVWRWTRTPPKPEYEPKRNLSRRVQALAA 188

Qy 217 --IKLKROKAVY-TAIVLPLSIVSLSVWTHVMSDN-----TNOVVPY----C 257
 Db 189 IIGLISLGHEHALYQVSAISLTSYTRRIQMCANITTPSFNNYMQNTNDYVFOPLPSPIAVL 248
 Qy 258 ILDNLNTAMGAWMFLICEAMSTAHLLAERFQALKHIGPAAWADYVTLWMLSLKJTRD 317
 Db 249 ILAVCTPFV----WNTMDLFIMMISKG-LSTPFEQ-----ITTRIKLHE 288
 Qy 318 TGNALCYTFVMSLVLFFITLSTIYGLMSQLSEGFGIKDIGLITALMV-IGLIFYCDE 376
 Db 289 --EVCS-----VPIQIREHIVKOCLE--FUDSAMSLILSVCNVNLVFC-- 332
 Qy 377 AHYASVNUVNTNFOCKLLMVLWLNW-----MSDAQTEIN 409
 Db 333 -YOLINN-----LRWPINIVYFWYSLVLYLIGRTAVFLTAADINEESKRGIG 381
 Qy 410 MFLRAT-----EMNPSPSTINCSCG--FEDVUNRTEFGKGLTTMVTYVULQF 452
 Db 382 VLRVVSSRSWCVEVERLIFQMTTQVALSGKCFPLTRRLFLFGMAGTIVTYELVLLQF 439

RESULT 15
 US-10-047-328-70
 ; Sequence 70, Application US/1047328
 ; Publication No. US20040003419A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Carlson, John R.
 ; APPLICANT: Clyne, Peter J.
 ; APPLICANT: Warr, Coral G.
 ; APPLICANT: Yale University
 ; TITLE OF INVENTION: No. US20040003419A1 Taste Receptors in Drosophila
 ; FILE REFERENCE: 44574-5072
 ; CURRENT APPLICATION NUMBER: US/10/447,328
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: US/09/593,519
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/138,668
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 60/181,704
 ; PRIOR FILING DATE: 2000-02-10
 ; NUMBER OF SEQ ID NOS: 94
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO: 70
 ; LENGTH: 477
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; US-10-047-328-70

Query Match 4.5%; Score 119; DB 15; Length 477;
 Best local Similarity 19.9%; Pred. No. 0.013; Gaps 22;
 Matches 101; Conservative 73; Mismatches 171; Indels 162; Gaps 22;
 Qy 108 RIIGTLP-IVRHGRPARAKFENNSAIFIYSTVPPVLLACTVGGYANRRIHV-----RSL 160
 Db 23 KIAGILPODLEKFRSRNLLEKSRNGMIV-----MLSTLILYVVLNLICLISFGBEDRSL 76
 Qy 161 SGPFEEAVIA--IPLVNLPIPMIPLWFEARKIAKLTNDMDFEVLYQI----- 210
 Db 77 KA-SOSTLTFVIGLFLTYGLIMMV-----SDQTLAKNQGRIGE-LYERIRLDERL 127
 Qy 211 --SGHSLPLKLRQKAVIATVPLIPLSVLSVI-TAVTMSD-----LNINQVVPYCI-- 258
 Db 128 YKEGCMNDNSTIGRIRIMLTVFELSTVLUVQYQSQMSLWVSAIPFINT 187
 Qy 259 IDNLTAMGAWMFLICEAMSTAHLLAERFQALKHIGPAAWADYVTLWMLSLKJTRD 310
 Db 188 LDKI-----WF-----AVSLYALKERFRAINATELVDTHEKHKLMLRGNOEVPP 234
 Qy 311 -----LSKLTRDTC-----NAJL-----Y 324
 Db 235 DSSOPPOYPSNLVEVLYKELGMDIGSIGKKSVEBKUNLCLQVHDBICEBIGKALNLSWY 294
 Qy 325 TVVFMSLYLFPPIITSIYGLMSQLSEGPGLKDIGITITALWNLIGLIFYCDBAHYASVW 384

Db 295 PIISLIMAYGLIFTAQLVPL-----YCATQYOSIPSLFSAKNRPFITVVL 340
 Qy 385 RTNFQKLUMLNEVNWNMSDAQT-----EYNMFLATEMAPST-INC 424
 Db 341 SYSGKCVLIVLIVSWMKTSQASKTGISHKCGUVADDNLYEVVNHSUKLHSVDSA 400
 Qy 425 GGFPDVNRLPKQHLLTAVTYVLUQFQISIPDKGSEGANNTVDFVMSLDNMS 484
 Db 401 CGFFTDMDTLYGVSGGITSYLLILIONNLAAQOKEAIQFN-----SLNTDAG 450
 Qy 485 LMGAST---LSTTT--VGTTLRPI 504
 Db 451 LVGATDMDNISSTLRDFVTTWMPAV 477

Search completed: May 23, 2005, 19:29:12
 Job time : 144 Secs

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein Search, using bw model

Run on:

May 23, 2005, 18:59:59 ; search time 41 seconds
 (without alignments)
 1201.536 million cell updates/sec

Title: US-10-081-816-12

Perfect score: 2645

Sequence: 1 MRPSCBKVKGHGQNSGHS.....TTVGTTLPPPIKMKRGK 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79;*

1: p1r1;*

2: p1r2;*

3: p1r3;*

4: p1r4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	119.5	4.5	C22845	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
2	118.5	4.5	T29461	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
3	112	4.2	F30010	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
4	111.5	4.2	395	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
5	107.5	4.1	346	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
6	107.5	4.1	H82509	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
7	106.5	4.0	583	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
8	106.5	4.0	698	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
9	105.5	4.0	702	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
10	105	4.0	453	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
11	105	4.0	A75145	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
12	105	4.0	B62425	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
13	105	4.0	B75059	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
14	104.5	4.0	896	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
15	104	3.9	369	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
16	103.5	3.9	317	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
17	103.5	3.9	T21047	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
18	103.5	3.9	T17358	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
19	103.5	3.9	515	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
20	102.5	3.9	D71127	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
21	102.5	3.9	A90255	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
22	102	3.9	468	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
23	102	3.9	A72096	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
24	102	3.9	925	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
25	102	3.9	E81573	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
26	102	3.9	D70117	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
27	101.5	3.8	1036	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
28	101.5	3.8	S34959	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
29	101.5	3.8	437	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
30	101.5	3.8	472	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
31	101.5	3.8	511	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
32	101.5	3.8	A97118	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
33	101	3.8	459	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
34	101	3.8	F90536	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
35	100.5	3.8	467	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
36	100.5	3.8	580	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
37	100.5	3.8	A83874	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
38	100	3.8	S42839	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
39	100	3.8	C41903	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
40	100	3.8	T14241	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
41	99.5	3.8	244	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
42	99.5	3.8	301	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
43	99.5	3.8	583	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
44	99	3.8	AH1151	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
45	99	3.7	360	NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion

ALIGNMENTS

hypothetical protein
 virulence factor m
 hypothetical protein
 NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
 thiazide-sensitive arsenical pump membrane protein
 NADH dehydrogenase (ubiquinone) (EC 1.6.9.3) chain 4 - Trypanosoma brucei mitochondrion
 probable permease glycerophosphoryl hypothetical protein
 melanocortin receptor

RESULTS

1

Query Match 4.5%; Score 119.5; DB 2; Length 439;
 best local similarity 16.4%; Pred. No. 0.055%;
 Matches 71; Conservative 82; Mismatches 127; Indels 153; Gaps 17;
 C;Species: mitochondrion Trypanosoma brucei
 C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
 C;accession: C22845
 R;Hengens, L.A.M.; Brakenhoff, J.; De Vries, B.F.; Sloof, P.; Tromp, M.C.; Van Boom, J.;
 Nucleic Acids Res. 12, 7327-7344, 1984
 A;Title: The sequence of the gene for cytochrome c oxidase subunit I, a frameshift containing a maxi-circle DNA.
 A;Reference number: A93537; MUID:85037915; PMID:6093040
 A;Accession: C22845
 A;Molecule type: DNA
 A;Residues: 1-439 <HEN>
 A;Cross-references: UNIPROT:Q33575
 C;Genetics:
 C;Genome: mitochondrion
 C;Superfamily: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
 C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation; C;Genetic code: SGC6

Query Match 4.5%; Score 119.5; DB 2; Length 439;
 best local similarity 16.4%; Pred. No. 0.055%;
 Matches 71; Conservative 82; Mismatches 127; Indels 153; Gaps 17;
 C;Species: mitochondrion Trypanosoma brucei
 C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
 C;accession: C22845
 QY 105 WFLRITIGVPIVRHGPARAKFEMNSAFTIYVVFVFLACIVGVYVANNRIRHIVSLSGP 164
 Db 51 WFVFFMGII-----MYLIFLSSKG-----VSVNYKPYI-----80
 165 EBAVIAVFLR-----VNLLPIMIPIILWYEARKIAKLFNDWDD-----FEVVYQISGHSLP 216
 Db 81 ---VMIKMYIYIYVNLVLIIL-----DDFCMCFIAFESLFFPCLVSLP 120
 QY 217 LKLRLKAV---YIAIVPILSVSVWVTHVMSD---LNIQNVVYPCYCLDNLYAMLGW 269
 Db 121 FNFNRTPFIAFLYVFLSSVYVCIICITVISHNINLQAPIDVYCFDLSAIFIW 180
 QY 270 WFLICEAMSTIAHLLAERFOKALKIGPAMVADYVRLML-----RUSKLR 316
 Db 181 ILLF-----IMPAKIPWPH-----VWLPPEMVAVTEMVSLAVIL 219
 317 DTGNALCYTFVMSL-----YLFFITISIYGM-----SOLSEGEKIGKIDG 358
 Db 220 KIGFFCVYKFLFIAFTNTISIWPLGFDISVIVLGLVFLAMSLLFSDYKLLIANWVHNG 279
 QY 359 LTITALWNLGILF-----YICDEAHYAVN-----VTFNFOKGLWYELWNWNSDAQ 406
 Db 280 IGLILWHDILFLVGLLILCILNHLSSRFMFIVIGMIDNYGVIFLILSFRGSIWS 339
 QY 407 E-----INMELRATENPSTINGGPFDFVRLKGELITWVYV 447
 Db 340 SLFLCILFLFENIDFPFLMLFLYVDFILYGLISIISFYIISPVYIITFLFSLI-----YV 394

RESULT 2

QY 448 VLLQPOQISIPDVK 460
 : ||| : |||
 DB 395 MCJSEFSFWLDK 407

T29461 hypothetical protein T05B11.2 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T29461
 R;Geibel, C.; Bradshaw, H.
 Submitted to the EMBL data library, March 1996
 A;Description: The sequence of *C. elegans* cosmid T05B11.
 A;Reference number: Z20621
 A;Accession: T29461
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-789 <GEN>
 A;Cross-references: UNIPROT:Q22211; EMBL:U53151; PIDN:AAB37067.1; GSPDB:GN00023; CESP:T05B11.2
 A;Experimental source: strain Bristol N2; clone T05B11
 C;Genetics:
 A;Gene: CESP:T05B11.2
 A;Intron: 2/3; 123/2; 229/2; 344/1; 362/1; 418/3; 512/2; 618/2; 707/2
 Query Match 4.2%; Score 112; DB 2; Length 443;
 Best Local Similarity 20.2%; Pred. No. 0; 23; Indels 128; Gaps 20;
 Matches 82; Conservative 65; Mismatches 131; Indels 128; Gaps 20;
 QY 97 YRNI-DPIN-WFLRITGIVLPIVRHGPAREKPEMNSASFIVSVFFVLLACYVGIVYANR 153
 Db 43 YNIVIYANVINYIWFIVYFNGLIVPF----LIPFLSRLKLVSKYKFYILL----- 85
 QY 154 IHIVRSLSGPBEEAVAYLFVNILPIMIDILWYEARAKIAKLFNDWD-----PEVL 206
 A;Cross-references: UNIPROT:Q22211; EMBL:U53151; PIDN:AAB37067.1; GSPDB:GN00023; CESP:T05B11.2
 A;Experimental source: strain Bristol N2; clone T05B11
 C;Genetics:
 A;Gene: CESP:T05B11.2
 A;Intron: 2/3; 123/2; 229/2; 344/1; 362/1; 418/3; 512/2; 618/2; 707/2
 Query Match 4.5%; Score 118.5; DB 2; Length 789;
 Best Local Similarity 20.8%; Pred. No. 0; 14; Indels 170; Gaps 20;
 Matches 100; Conservative 65; Mismatches 20; Indels 145; Gaps 20;
 QY 37 FLNAAKPL-----NSANAOQAYLGYRKVKSIGLAKERDADYEAPPLDRKGSDDSTASN 88
 Db 230 FPPNSNPLVLSIILSILNLAATIIFYVFLVFLPCTQDEHL-----KSTNCLINN 277
 QY 89 PEFKPI---PSVVF-YR---NIDPNWFLRIGIVLPIVRHGPARKFEMNSASFIVSVF 138
 Db 278 DEPECTKDPFLIVQTRIMIDRLQKETRICIIASPIQTQMFFACTFLINFGELLIK 337
 QY 139 FVLLACYVGIVANRRIHTVRSLSGPPEEAIVAYLFVNILPIMIDILWYEARAKLFN 198
 Db 338 KVIL-----ANEIDKLMYKMSN-----LPVFSFDDDKKSEBLG 370
 QY 199 DWDDPDEPVLYQIISGSLSPKURQKAVVATVLPISVLSVITRTMSDINTNQVPCY 258
 Db 371 PFRHF---LKVSLUDCSAKTKKCCSIFTCLV---GVIVLITFPRKISVLMQTAQPF 422
 QY 259 LDNITATMLG-AWMLICE-AMSTIAHLLAERFQKALKHIGPAMWA---DYLVML 309
 Db 423 SWGERSFFGYPAIKFVFCVSLCLFGHTKNGLASKFCKRLVVRVMILQVANPKLDFRILHI 482
 QY 310 RLSKLTDRDGNALCUTF---VFMSSLYLFITLISIY-GIMSQISEGFGKIDQIQTITA 363
 Db 483 L-----ALCHSIPWVAIMSWIIVPINKIYGGIEQNLISRIFLIVLWLYIWIY 532
 QY 364 LWNIGLFLYFICDEAHAYASVNRNTNQKCLMVELMNNSDQ----- 405
 Db 533 ISTICLAIYI-----PMSALNREVNIVFNEBLQAKEBKTLNIGVLEKFD 578
 QY 406 -----TEINMFLRATEMNPTINGGGFPVNRTFPGLL-----TTMVTYLVLV 449
 Db 579 FRQNFILEMVLF----ANGSLSSGGFAPL-FLWGLVNGIYIITSFFNTPVLYCIL 631

RESULT 3

F30010 NADH2 dehydrogenase (ubiquinone) (BC 1.6.5.3) chain 4 - *Leishmania tarentolae* mitochondrion
 C;Species: mitochondrial *Leishmania tarentolae*
 C;Date: 30-Jun-1999 #sequence_revision 30-Jun-1999 #text_change 03-Jun-2002
 C;Accession: F30010
 R;de la Cruz, V.F.; Neckelmann, N.; Simpson, L.
 J; Biol. Chem. 259, 15136-15147, 1984
 A;Title: Sequences of six genes and several open reading frames in the kinetoplast maxid

QY 163 PFEAAVIAFLVFLN-ILPIMTIPILWYE-----ARKIAKLFNDWD-----FE 204
 Db 41 PDKVVLVLSIINATILLITQKIGLISFFISLTOGFAQFLVFLVLSKFETQRULFN 100

RESULT 4

164248 hypothetical protein homolog MG443 - *Mycoplasma genitalium*
 C;Species: *Mycoplasma genitalium*
 C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C;Accession: I64248
 R;Frasier, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.; C.A.; Venter, J.C.; Science 270, 37-403, 1995
 A;Title: The minimal gene complement of *Mycoplasma genitalium*.
 A;Reference number: A64200; MUID:96026346; PMID:7569993
 A;Accession: I64248
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-95 <TIGR>
 A;Cross-references: UNIPROT:P47681; GB:U39731; NID:9345031; PIDN:AACT2463.1
 A;Experimental source: strain G-37
 C;Genetics:
 A;Gene: SG3
 C;Superfamily: Mycoplasma-pneumoniae hypothetical protein yf1b

Query Match 4.2%; Score 111.5; DB 2; Length 395;
 Best Local Similarity 20.9%; Pred. No. 0; 22; Gaps 19;
 Matches 78; Conservative 63; Mismatches 123; Indels 109; Gaps 19;

RESULT 5

H82509 hypothetical protein VCA0023 [imported] - *Vibrio cholerae* (strain N16961 Serogroup O1) C:Species: *Vibrio cholerae* C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004 C:Accession: H82509 C:Heidelberg: Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Brmolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. L.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. A:Reference number: A82035; MUID:20406833; PMID:10952301 A:Accession: H82509 A:Status: preliminary A:Residues: 1-583 <GLA> A:Cross-references: UNIPROT:Q92E36; GB:AL592022; PIDN:CAC95857.1; PID:916413065; GSPBB: A:Experimental source: strain Clip11262 C:Genetics: A:Gene: Lin0625

Query Match 4.1%; Score 107.5; DB 2; Length 583; Best Local Similarity 21.4%; Pred. No. 0.76; Gaps 24; Matches 84; Conservative 66; Mismatches 104; Indels 139; Gaps 24;

QY 106 PFLIGLPIRWH--GPARKRPEMSASRFTYSVVFVLLACYGVANNRRIHIVRSLSG 162

QY 18 YIKVYUULLTQIAOFVIGP-----FVYFFFILRVIGVPGITDANL-----G 59

QY 163 PFEAVAYTLEFLVNLIPIMIPIFLWTEARKIAKLFNDWDDDEVL--YQISGHSPL-- 217

QY 60 EVFSSPVAVIML-IAALLLFLVYEL-----GFTIMMAYQLRGESYVPKI 107

QY 218 --KLRGKA-----VIAVLPILSULSVVITHVMSDLINQVYPCYCLDNL- 262

QY 108 IORLYNTAKYVFLPSVQAYVFLYFLFLPIAG-LSLPIT-----ITENLYLPHFIDLM 160

QY 263 TMLGAWHFLCEAMSITAHLLAEROKAHLWIGPAMWYRVMWRLSKLTROTNAL 322

QY 161 KTTGTCWLVYI--AIAIYFSA-RIVFALLY-----FIBKS--LKVIS-----GAI 202

QY 323 CTFVFLSFLVYF-----ITVSIYG-----LMSQISBGFGJRKIDGIT 362

QY 203 RSKWFSQSKHIFPMLKWLIVWIGLVLISIATIMPLVPLVKTRGIAVIGILT 252

QY 363 ALWNIGL-----FVYCDAHAYSVNVRNP--EKKLMLWELNNNSD 403

QY 263 ILOVIGPFAAGIFQGITAQLVKAFAI--EQQHAPV-ARQQPFHKRKFITV----- 311

QY 404 AOTBIMPLRFLMENSTINGCGFPDRTRK 436

QY 312 --AIVVFIISGFN-----IFTVNATLYE 333

RESULT 6

Db 297 IWWIMMNN-----LAREGLAGLKDTEFLGFLUTTYLWKGIERTRPKL 344

Db 101 ILYW-----LLYVFNIPILITFSYKIGKNTIISTHFWA---SNVFGP 142

Db 143 LISTPGSDNPPM-----ASITTFWQAKDQJNQASGFVPP--LMSFT 185

QY 257 CI-----LDTNTAMGAWNPLICEAMSITAHLLAEROKAHLWIGPAMWYRVMWRLSKLTROTNAL 311

QY 312 SKUJRTDTGHALCYTFVPMFLYLFIT-ITSIYGMQSLQSGFGI-----RDIGLT 360

QY 187 SO-----GKVIISTPITVAAIYGFNGISVLLYIIGGSAAGDADFLTOYARKKRNNSV-- 239

QY 361 ITALWNIGLFLVYICDABAHASVNTNFQKLMELWNNSDQTEBIMFLRTEMIS 420

QY 240 -----SILFVYVNSFPLITAILIGSPVFGSLLQDQN-NYRDSAWEVSLF-----FSPN 286

QY 421 TINCGGFDVNRTRFLFKGLIT-TWVTVLWVLOF-QS1PTDKGPSSEGANNITVVFVMSD 478

QY 287 LI-----ATFFSILLRGTVVSVLFFPRINPAKVKFDK--LBEVKALIUSDNNAS 335

QY 479 LDNDMSLMGASTL 491

QY 336 LSTQETFLGGYVSL 348

RESULT 7

Db 95228 hypothetical Protein SP193 [imported] - *Streptococcus pneumoniae* (strain TIGR4)

C:Species: *Streptococcus pneumoniae* C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004 C:Accession: D5228

C:Accession: D5228

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heijnen, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, T.; Hickey, E.K.; Holt, T.E. Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, P.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: D95248
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-698 <KUR>
A;Cross-references: UNIPROT:Q97NR4; GB:AB005672; PIDN:AAK76021.1; PID:g14973459; GSPDB:G
A;Experimental source: strain TIGR4
A;Genetics:
A;Gene: SP1953

A-Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A-Reference number: A7872; MNUID:2122245; PMID:11544234
A-Accession: H98092
A-Status: preliminary
A-Molecule type: DNA
A-Residues: 1-702 <KUR>
A-Cross-references: UNIPROT:Q8DNF4; GB:AE007317; PIDN:AAL00573.1; PID:915459452; GSPDB:G
C-Genetics: -

Matches	Local Similarity	Score	DB	Length
4.0 %	19.1 %	Score 106.5 ; Pred. No. 1.1 ;	DB 2 ;	Length 698 ;
6	EVKVKHG-----GOONSCHLSGCMANYYRKKGDAVFLNAKPLNSANAQAYLGVRKYS	59		
88	EKISKKNKFYILDPAKGRQRQMSISEFERHYSN-----TILTFFKKLDS-----FMSRK--	133		
60	IGLAERLIDADYEAPPLDKKSSDSTASANNBFKPSVYFYRNIDPWNFLR-----IIGVLPL	114		
134	-----DNKAS-----PVLYKFYKRNKLGLFFFTALLYVQSLP	169		
115	IVRHGPARAKPEMNSAS-----FLYSVVF--FVLLACYVG-----	147		
170	IANRVIDTWFKDDSYISRMALFTLIFITFSFLSMYLRLQTYVASTIKYIMDEKISYDFMK	229		
148	-----"YVANNRIRHIVRSLSLGSPFEEAVIYL-----	172		
230	HLIYLPYSFYEKRTLGDLFRANSTVYIREISNNFAILDLMMIVVYAVLFSPSKYM	289		
173	--FLVNI--LPIMIPIPLWEARKAK-----	197	-LF	
290	VIFLSSLALSIWVYPIKISKNLIDKNTKEKVWQNITSEVLSKNSDIKLTGEEFWI	349		
198	NDWDDPEVLYVQIISGHSLPLKLQKAVVIAVLPILSLSVWVTHWM-----SDL	248		
350	NKWDNFNTKOLIG-----RKDHILSTVSSITVNLQITLPVILQVNIKTFEQ	401		
249	NINOVV-----PYCI-----LDN-LTAMLGAWMFL--ICRAMSITAHLAERFOKALH	294		
402	TLGQIVIASTVSPYFSPISPLISLSDNYIQMLMKYGPRLIEDVENTKSELIPERSODIKF	461		
295	IGPAAMADVRY-----LWLRSLKUTRDITGNALCYTFVFMSSYLFFFTLITSYGIMSQLE	350		
462	-----DKKIELKDIDWYK-----YGLP-----D	478		
351	GFGIKDIGHI-----TALWNITGLLFYICDEAHYASVNV-----	384		
479	DIVUKGGINVYIKKGFTVAIYGSGSKSTLAKIULGLL-----EPNGISTEDGVKEE	533		
385	-RTNFQKLUMLVLMWNNNSDAQTBIM-----PLRATEMNPSTINGCGFFDVNTRAFKGLL	439		
534	GQTLVYRKIFGAVLONSTLSYGTLENITFGHVFVSDEEL-WTNLNSTGLSNVVKSLPGL	592		
440	TMV-----TYLWLUQFQISIPTDKGDSSEGANNITWDYMDSDLNDMSLUMGAS	489		
593	TIATBEGNNFSGQGQOMILLARCLLSKP-----SVVWLDBATSSLNLSQOITTS	642		
490	TLSITVGT 499			
643	YLS-EIGHT			

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain L1 RP792 - *Rickettsia prowazekii*
 C;Species: *Rickettsia prowazekii*
 C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
 C;Accession: B71640
 C;Anderson, S. G. E.; Zomorodiour, A.; Andersson, J. O.; Sicheritz-Ponten, T.; Alsmark,
Nature 395, 133-140, 1998
 A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.
 A;Reference number: A71630; PMID:99039499; PMID:9833893

A;Molecule type: DNA	Db	166 PY----TII-----LEIFNPNSPFPRY-FHMITAS--VITTSFVIGGASFYLLNR 211
A;Residues: 1-653 <AND>	Qy	154 IHRVRLSGPPEEAVIAYLEFLVNLPIIMI-----IYVEARAKLENDW 201
A;Cross-references: UNIPROT:Q9ZCG1; GB:AJ235273; GB:AJ235269; NID:93861237; PIDN:CAA1521	Db	212 -----YKCHAKIMLFENAVMLALIVSPIOFIGDHLGIAINTLKVOPVKSATIGW 261
A;Experimental source: strain Madrid E	Qy	202 DPEVLVYQISSHSLPLKURQAVYIAVLPILSWSVWITW-----MSDLNINQVP 255
C;Genetics:	Db	262 TEKGASFLIG--LPDKEREKTY AIEIPAS- SLILTSLDSGEVKLGKETWKRPP 316
A;Gene: nub11; RP2162	Qy	256 YCIL-DNLTAALGAWFLICHEAMSTAHLLER-----FOKALKHTGPAWADYRVL- 307
C;Superfamily: membrane-associated complex; NAD; oxidoreductase	Db	317 VAVFVFSFRIMLIGGCLMVPFGIAGLYLYLNKLYTYWFWYIILMSPSRIA---VLA 373
Query Match 4.0%; Score 105.5; DB 2; Length 653;	Qy	308 -WL-----RUSKLTRDTGNALCYTFVPMISLYLFRITLISYGL 344
Best Local Similarity 20.3%; Pred. No. 1.3; Mismatches 95; Conservative 76; Mismatches 178; Indels 119; Gaps 24;	Db	374 GMLVTEVGROPVIVNLK-TVDTSPLIGKVFISLIAFWVWYLLIRGV 422
Qy	153 RIHTWLSLCPFEEAVIAVLEFLVNLPIIMIPLAEEARKIAKLENDWDDPEVLYQIISG 212	
Db	113 K-GIRFLS-----PLSLPFTFMLMLV-----SADNFLQLFCGWEVGVCYSLIG 157	
Qy	213 HSLPLKURQAVYIIVLPL-----ILSVLWSVWITW-----NINQVVP 255	
Db	158 FWHSGESANGAIAKIFRITNVSDFRPAKAFEMNSASFTYVVFVLLACYVSVANN 152	
Qy	218 FSILIDICLIL---FIGCMGRSAQJGLHNLWLPDAMEGPTVPSALIH--AAITWTRGVFL 271	
Db	308 WRLRSLKLTTRDTGNALCYTFVEMSLVLPFILSIVGMSQSEGCI--KQIG--LTIA 363	
Qy	272 VARCC-----YLFEVSPVQIOPITIG--GICLFLAASIAIMQSDIKKLTAYST 318	
Db	364 LWNIGLILFYICDEAHAVASV--NVRTRNFOKKLIMLNEWNNMSDQTEBIMELRATENPS 420	
Qy	319 CSQKLYMPMACGVSSINSALFHVLVTAFFKALFLPSAGNV-IHVNENIFRKGMLINKM 377	
Db	421 TINGC-----GFPDVNRTL-----FKGLLTMMVYL---VVL 449	
Qy	378 PITYGNFLICSLALGITYPLSGPSKQDILEATYSSGFSMFTFGITTAILTAYSMKII 437	
Db	450 LQFQTSIPTKGDSGEGANITVDFMDSLDMMSLMSGASLTITVG 477	
Qy	438 LVFHFKTKEKDVFVKAHEPKI-----MONPLILVAGSPFSGMIG 479	
RESULT 10		
H7132 cytochrome D ubiquinol oxidase chain 1 (cyda) RP216 - Rickettsia prowazekii	Db	167 IHRVRLSGPPEEAVIAYLEFLVNLPIIMI-----IYVEARAKLENDW 201
C;Species: Rickettsia prowazekii	Qy	212 -----YKCHAKIMLFENAVMLALIVSPIOFIGDHLGIAINTLKVOPVKSATIGW 261
C;Accession: H7132-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004	Db	202 DPEVLVYQISSHSLPLKURQAVYIAVLPILSWSVWITW-----MSDLNINQVP 255
R;Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Aisemark, T.	Qy	262 TEKGASFLIG--LPDKEREKTY AIEIPAS- SLILTSLDSGEVKLGKETWKRPP 316
Nature 396, 131-140, 1998	Db	256 YCIL-DNLTAALGAWFLICHEAMSTAHLLER-----FOKALKHTGPAWADYRVL- 307
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.	Qy	317 VAVFVFSFRIMLIGGCLMVPFGIAGLYLYLNKLYTYWFWYIILMSPSRIA---VLA 373
A;Reference number: A71630; MUID:99039499; PMID:9823893	Db	308 -WL-----RUSKLTRDTGNALCYTFVPMISLYLFRITLISYGL 344
A;Accession: H7132	Qy	374 GMLVTEVGROPVIVNLK-TVDTSPLIGKVFISLIAFWVWYLLIRGV 422
A;Status: preliminary; nucleic acid sequence not shown; translation not shown	Db	166 PY----TII-----LEIFNPNSPFPRY-FHMITAS--VITTSFVIGGASFYLLNR 211
A;Molecule type: DNA	Db	154 IHRVRLSGPPEEAVIAYLEFLVNLPIIMI-----IYVEARAKLENDW 201
A;Residues: 1-453 <AND>	Qy	212 -----YKCHAKIMLFENAVMLALIVSPIOFIGDHLGIAINTLKVOPVKSATIGW 261
A;Cross-references: UNIPROT:Q9ZDV3; GB:AJ235270; GB:AJ235269; NID:93860572; PIDN:CAA1467	Db	202 DPEVLVYQISSHSLPLKURQAVYIAVLPILSWSVWITW-----MSDLNINQVP 255
A;Experimental source: strain Madrid E	Qy	262 TEKGASFLIG--LPDKEREKTY AIEIPAS- SLILTSLDSGEVKLGKETWKRPP 316
C;Genetics:	Db	256 YCIL-DNLTAALGAWFLICHEAMSTAHLLER-----FOKALKHTGPAWADYRVL- 307
C;Gene: cyda; RP216	Qy	317 VAVFVFSFRIMLIGGCLMVPFGIAGLYLYLNKLYTYWFWYIILMSPSRIA---VLA 373
C;Superfamily: cytochrome d complex terminal oxidase chain 1	Db	308 -WL-----RUSKLTRDTGNALCYTFVPMISLYLFRITLISYGL 344
Qy	440 VGLYLYRFYEE 449	

Qy	96	FYRNIDPIMWFLRTIGVLTIVRHG--PARAKFEMNSASFLTYSVVFVLLACYVGVANNR	153						
Query Match	4.0%	Score 105; DB 2; Length 453;							
Best Local Similarity	23.4%	Pred. No. 0-89							
Matches	68;	Conservative	50;	Mismatches	98;	Indels	74;	Gaps	17;

RESULT 12
B64245
periplasmic phosphate permease homolog AG88 homolog - *Mycoplasma genitalium*
C;Species: *Mycoplasma genitalium*

C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
 C;Accession: E62425
 R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
 M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.;
 C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A;Title: The minimal gene complement of *Mycoplasma genitalium*.
 A;Reference number: A64200; MUID:96026346; PMID:7569993
 A;Accession: E62425
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-654 <TIGR>
 A;Cross-references: UNIPROT:P47651; GB:U39726; GB:L43967; NID:gi046113; PID:gi046125; TI
 A;Experimental source: strain G-37
 C;Genetics:
 A;Genetic code: SGC3
 A;Start codon: GTG
 C;Superfamily: periplasmic phosphate permease AG88

Query Match 4.0%; Score 105; DB 2; Length 766;
 Best Local Similarity 20.1%; Pred. No. 1.7; Gaps 24;
 Matches 111; Conservative 70; Mismatches 188; Indels 182; Gaps 24;
 Qy 89 PFKPSVYRNIDPIN----WFLRITGVLPVTRHGPARKETEMNSASFIYSSVVFVLLA 143
 Db 46 PDKAKSLFNLNEFLGNKQAGIWP----PLI----VSPIVS-IGALIA 84
 Qy 144 CTVG-----YVANRH-----IVRSLSQPFPEAVIAYFLVNLNIPMIPILWYE 190
 Db 85 SIVGVRSTSFLYVTRCKRKIRKKGSLDILS--IPSIVFGIFASQIISIPF----- 134
 Qy 191 RKLAKLFLNDWMDPFEVLYQISCHSLPLKLRQAVIA-IVRL-----SVL 236
 Db 135 RDIKL-----PLSILVNTIAMSFMIPIVLSITNTLYVNDL 176
 Qy 237 SVIHTHMTSDIININQVPCYCLDNLNTAMLG-AWMLICEAMSIATAHLAERFQALKH 295
 Db 177 SWSVUGENKTSIAKIKKEKPQJWVLTAFARAISETMWNFLQSQSYNQEVNN- 235
 Qy 296 GFAAMVADYRVWLRISK-LTRDTG---NAACYTF---VEMSLYLFITI-----TL 339
 Db 236 -NRPFTSLDKITGGSVSTFIFSENGDQBQINGULYFGIILILVSLANFAWNSAPKTL 294
 Qy 340 SIVGILMSQLSBCFGKIKDQIGLTITALW---NIGLFYVCDERAYA-----SVNRTNF 388
 Db 295 ERYPEFLKKS-----NFTIQVWMPNNSALFVDTSTRQSVKRKYNNIERSLF 346
 Qy 389 QKLLM---VELNNWNSDAQPEINMLR-----ATEMFSTINCGF 427
 Db 347 FFRERQSVWVWIKLNYFELKIFOBELCTELAFGEVLAIIFLVFLNGSVAINNNGSTVES-- 403
 Qy 428 FDVNRLTFLKGILTTMVLQYFLQFQISIP-----TDKGDSERGANNITVDFMDS 478
 Db 404 FEADST-GRALVNTLV--ILITITPPLALLIAIWNTNTYNSKVKVN-VFNNTVIDS 457
 Qy 479 LNDMSLM 486
 Db 458 LSMPSTI 465

RESULT 13

B75059 probable transmembrane oligosaccharyl transferase PAB0974 - *Pyrococcus abyssi* (strain Or)
 C;Species: Pyrococcus abyssi
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C;Accession: B75059
 R;Anonymous, Genoscope submitted to the EMBL Data Library, July 1999
 A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A;Reference number: A75001
 A;Accession: B75059
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-766 <KAW>
 A;Cross-references: UNIPROT:Q9GP55; GB:AJ248287; GB:AL096836; NID:95458657; PIDN:CAB5036;
 A;Experimental source: strain Orsay
 C;Genetics:
 A;Gene: PAB0974

Query Match 4.0%; Score 105; DB 2; Length 766;
 Best Local Similarity 20.1%; Pred. No. 1.7; Gaps 24;
 Matches 111; Conservative 70; Mismatches 188; Indels 182; Gaps 24;
 Qy 18 GHISLGMANYTRKKGDAVFLNAKPLNSANQAYLYGRKYSIGLARLD----ADYEA 72
 Db 139 GHIFRSVHANVIR--GD-----NIVMLFWTSLLAGIAYLRTKQRFLVYLV 183
 Qy 73 PPDURKKSSDSTASNNEPKPSVYRNIDPINWFLRITGV-LPTIVRHPARAKFENNS 131
 Db 184 PTLASGTS-----SVFWQAYPLFVFSLINGVFLAIGSFLDKRKNFLDSFI 230
 Qy 132 FLYSVVFFVLLACV-----GYVANRRIHVSISGPFEAEVATLF--IUNLIPIM-- 181
 Db 231 IILSTAFGATIANYLGEFKFGYGMGLGYNRQHIVSKLGIKGKIRDAYLFTIHLVLPISLG 290
 Qy 182 IIPILWEAR-----KIAKL----- 196
 Db 291 LILITLFLSRFVTSKSKAAGVGVIGLTVSILJILKPAFLRGLGJFDMFKSTPIMETRP 350
 Qy 197 --FND-WDDEFV-----LY----- 208
 Db 351 TNFHDLMKAFSISIFLPLFLRFPBKVKTEDFFLGLIVPSLMLAWARFVFGSLA 410
 Qy 209 -----QISCHSLPLKLR-----KAVYIATVPLISVLSVWTHVMSDL----NINQV 253
 Db 411 VATMAGIGLVLGVSYLVIQRRKSGASRVALILIL--LIVINGAFTLKLCSMRPLINKE 468
 Qy 254 VPYCIL-----DLNTAMLGAWMFL-----ICEAMSTAHHLAERFQALKH 293
 Db 469 WENALIWLKNNSENNDVILAWMDYGAWITYSRRAPVATAPNPDVALLYGARNRDMW 528
 Qy 294 HIGPAAMVADYRVWLRISK-LTRDTGNACTFVEMSLYLFITIISIYGLMSQISERGG 353
 Db 529 SLGQDVIVSY-YDFLKFESSIVRTASAQSGYNLKRYWIAVPLTSSGGIL--FEGGE 585
 Qy 354 KDGIGLTATWNLGFLYFCDBARYASVNTRTNFQKELKJAKVLENKNNNSDQTEM-FL 412
 Db 586 YKIAKP-GDIDWVRIT--IGDHVVYPR-GLIVEYKGKVISLKLKSYSTDALYINLYK 641
 Qy 413 RATEMPSTIN 423
 Db 642 YALMNESETFN 652

RESULT 14

H86762 ABC transporter permease protein ylbB [imported] - *Lactococcus lactis* subsp. *lactis* (strain Or)
 C;Species: Lactococcus lactis subsp. lactis
 C;Date: 22-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 C;Accession: H86762
 R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaiillon, O.; Malarrie, K.; Weissenbach, J.; Ehrlic
 A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp
 A;Reference number: A86625; MUID:21235186; PMID:11337471
 A;Accession: H86762
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-896 <STO>
 A;Cross-references: UNIPROT:Q9CQJ2; GB:AE005176; PID:912724061; PIDN:AAK05202.1; GSPDB:G
 A;Experimental source: strain IL1403
 C;Genetics:
 A;Gene: ylbB

Query Match 4.0%; Score 104.5; DB 2; Length 896;
 Best Local Similarity 22.3%; Pred. No. 2.3; Gaps 13;
 Matches 63; Conservative 42; Mismatches 111; Indels 67; Gap 13;

QY	157 VRSISGGPFPBAAVIAVFLVNLIPIM-----IIPIMWEKIAKLFNDW---D	Db	308 WILRSKLTDRD-----GNAICYPVFF
Db	: : : : : : : : : : : : :	Db	: : : : : : : : : : : : :
QY	361 IDSNISNPVTFALALLYSLTMTMRFVEERGNGLKLGYSNRDIRKFKFVWGLVSS	Db	204 GLKIFPKDLSDTLAVYFOQOLICLPGV
QY	202 DFEVLYYQTSGHSLPLKLROKAVYIAVLPILSVWVTHVMSDLNIN---QVVPY	Db	362 TALWNIGLFLYICOBAYASVNRVTRNF
QY	421 GLGALVGTIGTGF-----LFI-AVFAVYASSTFSNIRLITISPLWIVAF	Db	248 ----LGIWMAITWILIPY-----
QY	257 CILNLTAMIGAWNLLICEAMSITAHLLAERFQKALKHIGAAAMWVYRVLRSKLTR	Db	422 INCGGFFDVNTLTFKGLLTTMVTLYVW
QY	466 AIA-TACSLIPAYWVTRMELKEVPAASLFLAKVPA---GSRILRKINFWKRMSPFTY	Db	272 I--VGFM-----SPAKALFAIIVVIVVQK
QY	317 DTGNALCYTFVPMSCYLFFITLSTYGLMSOLSEGRGKQDGTITALWN--IGLIPYIC	Db	482 DMSLWGMASTLSTTWTGTLPPPI
QY	521 VTARNL--FRYKQMLMTI--FGAGCTTALVWMSGFIRD---SISGLSNKQFGQIL---	Db	318 -LLLIGAGSFG-GIGMILAVPV
QY	375 DEAHYASVNRVTRNFORKLMLVTELNNNNDACTEINFLRATM	Db	201 417
QY	570 ---HY-----DMTIEKNKVNDEKEEIDKEIASBI	Db	598
RESULT 15			
G0007	conserved hypothetical protein yueF - <i>Bacillus subtilis</i>		
Species: <i>Bacillus subtilis</i>			
C;Accession: G70007	#sequence_revision 10-Sep-1999 #text_change 09-Jul-2004		
C;Date: 10-Sep-1999			
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteletti, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chodat, P.; Ehrlich, S.D.; Emmerlein, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, B.; Galler, Nature 390, 249-255, 1997			
A;Authors: Foulonge, D.; Pritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, J.; Harwood, C.R.; Hensut, A.; Hibert, H.; Holsappel, S.; Hobson, S.; Hulio, M.F.; Koetter, P.; Koenigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Nasuda, S.; Mauel, Y.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivotta, C.; Rocha, B.; Roche, B.; Rae, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Schleicher, S.; Schröter, R.; Scoffreto, P.; Sekiguchi, J.; Sekowska, A.; Seror, A.; Auchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tohato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.			
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Nasuda, S.; Mauel, Y.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivotta, C.; Rocha, B.; Roche, B.; Rae, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Schleicher, S.; Schröter, R.; Scoffreto, P.; Sekiguchi, J.; Sekowska, A.; Seror, A.; Auchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tohato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.			
A;Title: The complete genome sequence of the Gram-positive bacterium <i>Bacillus subtilis</i> .			
A;Reference number: A69580; PMID:9804403; PMID:9384377			
A;Accession: G70007			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-369 <UNK>			
A;Cross-references: UNIPROT:032095; GB:299120; PID:92635613; RTDN:CA815168.			
A;Experimental source: strain 168			
C;Gene: yueF			
C;Superfamily: <i>Bacillus subtilis</i> conserved hypothetical protein yueF			
Query Match	3.9%; Score 104; DB 1; Length 369;		
Best Local Similarity	18.7%; Pred. No. 0.83; DB 1; Length 369;		
Matches	83; Conservative 65; Mismatches 139; Indels 156; Gaps 19;		
QY	105 WFLRIGLVLPIVRHGPAPAKFEMNSASRPFY-----SVWFTFLACVYGVYANRI--- 154	Db	155 --HIVRSLSGGPFPBAAVIAVFLVNLIPIM---IIPIMWEKIAKLFNDW---E 204
QY	9 WTLQQLFLVLL-----FVATKVSFVQFQPIVFLSTFLPMLAGILYVTFNPVTRL 60	Db	61 LEKKKRTLS---ILIVLIFGLAFASVGPIL--TAQVTLGFLFNNLDPDVKIQ 112
QY	205 VLYYQTSGHSLPLKLROKAVYTA-----VLPILSVLSVWVTHVMSD 247	Db	113 ALTKDLS-HSQWFTMMQDYVSIKSQSLTSPLQNLQPNQNTTSLSAVFGVNTITVI 171
QY	248 LNINQWVPCYICDNLNTMLGAMWFLICEAMSITAHLLAERFORKALKHIGPAAMWVYVL 307	Db	172 ITVPEFLPYMLKD-----GHRF---PHIAVUKLFASTYE 203

This Page Blank (uspto)

"The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195 (2000).
 RN [12]
 RP GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitfield E.D., Bayraktaroglu L., Bernan B.P., Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review.;"
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
 RN [3]
 RP IDENTIFICATION.
 RX MEDLINE=2107712; PubMed=11516643; DOI=10.1016/S0960-9822(01)00258-5;
 RA Dunipace L., Meister S., McNealy C., Amrin H.;
 RT "Spatially restricted expression of candidate taste receptors in the *Drosophila* gustatory system.;"
 RL Curr. Biol. 11:822-825 (2001).
 CC -- FUNCTION: Probable role in the gustatory response.
 CC -- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -- SIMILARITY: Belongs to the G-protein coupled receptor Dr-tr family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AE003479; AAF47803.1; --.
 DR FlyBase; FBrf005468; Gr65a.
 DR GO; GO:0016021; C:integral to membrane; NAS.
 DR GO; GO:0005027; F:taste receptor activity; NAS.
 DR GO; GO:005912; P:taste perception; sensory transduction o. . . ; NAS.
 KW G-protein coupled receptor; Glycoprotein; Hypothetical protein;
 KW Multigene family; Receptor; Transmembrane.
 PT DOMAIN 1 129 Extracellular (Potential).
 PT TRANSMEM 130 150 1 (Potential).
 PT DOMAIN 151 166 Cytoplasmic (Potential).
 PT TRANSMEM 157 187 2 (Potential).
 PT DOMAIN 188 222 Extracellular (Potential).
 PT TRANSMEM 223 243 3 (Potential).
 PT DOMAIN 252 252 Cytoplasmic (Potential).
 PT TRANSMEM 253 273 4 (Potential).
 PT DOMAIN 274 324 Extracellular (Potential).
 PT TRANSMEM 325 345 5 (Potential).
 PT DOMAIN 346 350 Cytoplasmic (Potential).
 PT TRANSMEM 351 371 6 (Potential).
 PT DOMAIN 372 433 Extracellular (Potential).
 PT TRANSMEM 437 457 7 (Potential).
 PT DOMAIN 458 512 Cytoplasmic (Potential).
 PT CARBOYD 431 531 N-linked (GlcNAc-. . .) (Potential).
 SQ SEQUENCE 512 AA; 57465 MW; B98BD94D619BFFF6 CRC64;
 Query Match 100.0%; Score 2645; DB 1; Length 512;
 Best Local Similarity 100.0%; Pred. No. 5e-189; Matches 512; Conservative 0; Indels 0; Gaps 0;
 Ov 1 MPPSGEKVKKGIGQGNSGHSLSGMANTYRKKGDAVFLNKAELNSANAOQVYIGVRKSYI 60
 Db 1 MPPSGEKVKKGIGQGNSGHSLSGMANTYRKKGDAVFLNKAELNSANAOQVYIGVRKSYI 60
 Ov 61 GIAERLADAYEAPPLDRKSSSTASNNPEFKPSVFRNIDDEINWFLRIGVLPVTRHGP 120
 Db 61 GIAERLADAYEAPPLDRKSSSTASNNPEFKPSVFRNIDDEINWFLRIGVLPVTRHGP 120
 Ov 121 ARAKFENNSASRISRYSVVFVLLACYVQVYANRRIHIVRSLSGPFERAVIAVFLVNLPI 180
 Db 121 ARAKFENNSASRISRYSVVFVLLACYVQVYANRRIHIVRSLSGPFERAVIAVFLVNLPI 180
 Ov 181 MIPILWYEAARKLFLNDWDPEVLYQIISCHSLFLKLRKAVYIAVFLVNLPI 240
 Db 181 MIPILWYEAARKLFLNDWDPEVLYQIISCHSLFLKLRKAVYIAVFLVNLPI 240
 Ov 241 TAVTMSDLNINQWVPCYLQDNLNTAMIGAWAWLICBMSITAHLLAERFKALKHGPAM 300
 Db 241 TAVTMSDLNINQWVPCYLQDNLNTAMIGAWAWLICBMSITAHLLAERFKALKHGPAM 300
 Ov 301 VADYRVVMLRSLKLTDTGNALEYTFFMSLFLFITTSLTYSGLMSQSLBFGKIDGLT 360
 Db 301 VADYRVVMLRSLKLTDTGNALEYTFFMSLFLFITTSLTYSGLMSQSLBFGKIDGLT 360
 Ov 361 ITALWNLIGLFLVTCDEAHVASYNVRNQKCLMVELNWNMSDAQTBRINMLRATENPS 420
 Db 361 ITALWNLIGLFLVTCDEAHVASYNVRNQKCLMVELNWNMSDAQTBRINMLRATENPS 420
 Ov 421 TINCGGFFDVNRLPKGLLTWVYLVLUQFOISIPTKDQSEGANNITWVDFVMSLD 480
 Db 421 TINCGGFFDVNRLPKGLLTWVYLVLUQFOISIPTKDQSEGANNITWVDFVMSLD 480
 Ov 481 NDMSLMGASTLSTTGTGTLPPIMKGRKG 512
 Db 481 NDMSLMGASTLSTTGTGTLPPIMKGRKG 512
 Ov 481 NDMSLMGASTLSTTGTGTLPPIMKGRKG 512
 Db 481 NDMSLMGASTLSTTGTGTLPPIMKGRKG 512
 RESULT 2
 O7PYF4
 ID O7PYF4
 AC O7PYF4
 PRELIMINARY; PRT; 379 AA.
 DT 01-MAR-2004 (T-EMBL; 26, Created)
 DT 01-MAR-2004 (T-EMBL; 26, Last sequence update)
 DT 01-MAR-2004 (T-EMBL; 26, Last annotation update)
 DE AGCP12374 (Fragment).
 GN Name=agCG5552; ORFNames=ENSANG00000017642;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_Taxid:180545;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAB01008987; EAA01019.1; --.
 FT NON_TER 1
 SQ SEQUENCE 379 AA; 42958 MW; 135438038754E0F2 CRC64;
 Query Match 48.2%; Score 1276; DB 2; Length 379;
 Matches 241; Conservative 50; Mismatches 72; Indels 0; Gaps 0;
 Ov 97 YRNIDPWNFIRIGVLPVTHGPARAKFENNSASRISRYSVVFVLLACYVQVYANRRIH 156
 Db 97 YRNIDPWNFIRIGVLPVTHGPARAKFENNSASRISRYSVVFVLLACYVQVYANRRIH 156
 Ov 4 FSSKPYLVRAGVLPVTRLPSGSGTAFVFLASPSMTCVLFLLTIVTAYFLNRIE 63
 Db 4 FSSKPYLVRAGVLPVTRLPSGSGTAFVFLASPSMTCVLFLLTIVTAYFLNRIE 63
 Ov 157 VRSLSGPFFEEAVAYFLVNLIPIMIPLWYEAARKLFLNDWDPEVLYQIISCHSLP 216
 Db 157 VRSLSGPFFEEAVAYFLVNLIPIMIPLWYEAARKLFLNDWDPEVLYQIISCHSLP 216
 Ov 64 VRTLGSRFEESVAVIYFLVNLIPIMIPLWYEAARKLFLNDWDPEVLYQIISCHSLP 213
 Db 64 VRTLGSRFEESVAVIYFLVNLIPIMIPLWYEAARKLFLNDWDPEVLYQIISCHSLP 213
 Ov 217 LKQKQKVVYIAVFLVNLIPIMIPLWYEAARKLFLNDWDPEVLYQIISCHSLP 276
 Db 217 LKQKQKVVYIAVFLVNLIPIMIPLWYEAARKLFLNDWDPEVLYQIISCHSLP 276
 Ov 277 MSITAHLLAERFKALKHGPAM 336
 Db 277 MSITAHLLAERFKALKHGPAM 336
 Ov 184 LSITAKILDAEDPORAHLRVGPAKVSSEYRSWLRLSKLARDTGFSTCYTIFICLYLIFI 243
 Db 184 LSITAKILDAEDPORAHLRVGPAKVSSEYRSWLRLSKLARDTGFSTCYTIFICLYLIFI 243

Query Match		Score	DB 1;	Length	447;
Best Local Similarity		26.1%;	Pred.	No. 46-229;	
Matches		115;	Conservative	89;	Mismatches 190;
Qy	Db				
PLNSANAOVLYGURKY--SIGLAERLADAEYAPPLDKKSSDSTASNNPEFKPSVFRN	42	15.3%;	Score 405;	DB 1;	Length 447;
PMLPNPQRQFLEDBVRVREKLUIMARGIAAMEEVVRKOBTVDDPLED----KHDSEYQT	18	100	IDPINWFLRIGVLPVIRHGP---ARAKPEMNSASIFTYVVFPVLLACYVGVANNI-	99	99
TKSLVLQFOIMGMPRHNNPERKLPRGTYSGWSKQVWMAIFIYSCQTTIVVULRERVK	74	100	100	100	100
HTYRSLSGPFEAEVIAVYFLVNLTPMIPIL-WYEAARKTIAKLFNNDPDEVLYQTSCH	155	100	100	100	100
KFVISPDRKFRDEALYNVIFISLIFTNFLIPVAVSRHGPQAVTFKPNMMWNYQKFKTRGS	134	100	100	100	100
SLPIKLQKAVIAVLPILSVIS--WVITHTMDSLNINQVYPCILDN-----	214	100	100	100	100
PIVENFLYPLTWSLCVFSWLSIAINTSO--YFLQDPDRWYTFAY	194	100	100	100	100
--LTAMIGAW--WFLICBAMSITAHLIILEROKAKLKHGPAAMVADYVULWRLSKLUR	262	100	100	100	100
YPIIAMNCFCSLWYINGNAFGTASRALSDALOPTTIGERKPAQGLTEVRLWDLHMQ	238	100	100	100	100
DTSNHALCYTFVENSLSLYFIFITISIYGMQ-LSEGEGKIGDITLTALWNQGILFYCD	317	100	100	100	100
OLGRAYSNMIGMCLVTFITFTIATYGSISEIDHGATYKEVGLFVIVFCMGLLYICN	298	100	100	100	100
EATVASYWVTRINFOKKLMLVNLWMMNSAQTENMFRATEMPSITINGGFEDVNRULF	376	100	100	100	100
EARTYASRKGVLDFQTKLNLNITAVDAAATOKEVEVMLVAINKNPPINMLDGYANINRBLI	358	100	100	100	100
KGLLTTMVTYLVLQLQFQTS	436	100	100	100	100
TTNISPMATYLVLLQFKIT	418	100	100	100	100
RESULT 6					
G9BD_DROME	ID_G9BD	DROME	STANDARD;	PRT;	412 AA.
AC	Q81M5;				
DT	05-JUL-2004	(Rel. 44, Created)			

DT	05-JUL-2004 (Rel. 44, Last sequence update)	CC	family. Subfamily I.
DT	25-JAN-2005 (Rel. 46, Last annotation update)	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
DE	Putative gustatory receptor 98d.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
GN	Name=Gr98d; Synonyms=Gr98B.1; ORFNames=CG31061;	CC	the European Bioinformatics Institute. There are no restrictions on its
OS	Drosophila melanogaster (Fruit fly).	CC	use by non-profit institutions as long as its content is in no way
OC	Endopterygota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	CC	modified and this statement is not removed. Usage by and for commercial
OC	Ephydriidae; Drosophilidae; Drosophila.	CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
OX	NCBI_TAXID=7227;	CC	-----
RN	[1]	CC	-----
RP	SOURCE FROM N.A.	CC	-----
RC	STRAIN=Berkeley;	DR	EMBL; AR003761; AAC14112.1; --
RX	MRIDLINE=201606; PubMed=10731132; DOI=10.1126/science.287.5461.2185;	DR	FBem046885; Gr98d
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	DR	G-protein-coupled receptor; Glycoprotein; Hypothetical protein;
RA	Ammanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,	DR	KW Multigene family; Receptor; Transmembrane.
RA	George R.A., Lewis S., Richards S., Ashburner M., Henderson S.N.,	FT	DOMAIN 1 7 I (Potential).
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	FT	TRANSMEM 8 28 1 (Potential).
RA	Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champs M., Pfeiffer B.D.,	FT	TRANSMEM 29 40 2 (Potential).
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	FT	TRANSMEM 41 61 3 (Potential).
RA	Abrial J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	FT	TRANSMEM 62 148 4 (Potential).
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,	FT	TRANSMEM 149 169 5 (Potential).
RA	Beezon K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	FT	TRANSMEM 170 181 6 (Potential).
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	FT	TRANSMEM 182 202 7 (Potential).
RA	Burkis K.C., Busam D.A., Butler H., Cadieu B., Center A., Chandra I.,	FT	TRANSMEM 203 244 8 (Potential).
RA	Cherry J.M., Cawley S., Dahake C., Davenport L.B., Davies P.,	FT	TRANSMEM 245 265 9 (Potential).
RA	de Pablo B., Delchev A., Deng Z., Maya A.D., Dew I., Dietz S.M.,	FT	TRANSMEM 266 294 10 (Potential).
RA	Dobson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	FT	TRANSMEM 295 315 11 (Potential).
RA	Durbin K.J., Evangelista C.C., Farraz C., Ferreira S., Fleischmann W.,	FT	TRANSMEM 316 369 12 (Potential).
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,	FT	TRANSMEM 370 390 13 (Potential).
RA	Grodeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	FT	TRANSMEM 391 412 14 (Potential).
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,	FT	CARBOHYD 127 127 15 N-linked (GlcNAc: .) (Potential).
RA	Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	FT	CARBOHYD 222 222 16 N-linked (GlcNAc: .) (Potential).
RA	Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	FT	CARBOHYD 412 AA; 47722 MW; 7B05DAF4FAC4D72 CRC64;
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	FT	-----
RA	Lasko P., Lei Y., Levinsky A.A., Li J.H., Li Z., Liang Y., Lin X.,	FT	-----
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	FT	-----
RA	Merkulov G.V., Milišina N.V., Mobarry C., Morris J., Moshrefi A.,	FT	-----
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzyk D.M., Nelson D.L.,	FT	-----
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	FT	-----
RA	Palazzolo M., Pittman G.S., Paon S., Pollard J., Puri V., Reese M.G.,	FT	-----
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	FT	-----
RA	Shieh B.C., Siden-Kiamois I., Simpson M., Skupski M.P., Smith T.,	FT	-----
RA	Spter E., Spradling A.C., Stimpleton M., Strong R., Sun E.,	FT	-----
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	FT	-----
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,	FT	-----
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	FT	-----
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	FT	-----
RA	Zheng X.-H., Zhong W., Zhou X., Zhu X., Zhu S., Smith H.O.,	FT	-----
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,	FT	-----
RT	"The genome sequence of <i>Drosophila melanogaster</i> .";	FT	-----
RL	"The genome sequence of <i>Drosophila melanogaster</i> .";	FT	-----
RN	[2]	FT	-----
RP	GENOME REANNOTATION	FT	-----
RX	MEDLINE=22426069; PubMed=12537572;	FT	-----
RA	Misra S., Crooby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,	FT	-----
RA	Hrdlicky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,	FT	-----
RA	Smith C.P., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,	FT	-----
RA	Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,	FT	-----
RA	Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,	FT	-----
RA	Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,	FT	-----
RA	Lewis S.B.;	FT	-----
RT	"Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review.";	FT	-----
RT	Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).	FT	-----
RN	[3]	FT	-----
RP	IDENTIFICATION	FT	-----
RX	Medline=2017560; PubMed=10710312; DOI=10.1126/science.287.5459.1830;	FT	-----
RA	Clyne P.J., Warz C.G., Carlson J.R.,	FT	-----
RT	"Candidate taste receptors in <i>Drosophila</i> .";	FT	-----
RL	Science 287:1830-1834 (2000).	FT	-----
CC	-- FUNCTION: Probable role in the gustatory response.	FT	-----
-- SUBCELLULAR LOCATION: Integral membrane protein (potential).	FT	-----	
-- SIMILARITY: Belongs to the G-protein coupled receptor Dr-tr	FT	-----	

Db 39 LGVVPYEFH-PATL----SLQFRSLILYCFCIAITIGPMGRCWVYLTPPTMNGVSA 92
 Qy 170 AY---LFLVNILPIMIP-ILWTEARKIAKLFNDWDDPDEBVLVQIISGHSLPLKLGKAV 224
 Db 93 HYTAMPMEILLDAILLILVPPCYLGRKKKGKLGFLU--VKKVYSSLISIFPURLWFRAC 150
 Qy 225 YIAIVLPLISLVSVTHVMDLANTINQVLP-----YCTLDNTLML---- 266
 Db 151 PL-----ENLYVDSAYGVNVNALVLPGIGIHEKVKKLISWSVISQLTAKIVLII 201
 Qy 267 -GAWFLFICBAMSSTAHLLAER----POKALK--HICPAAMWADYRVIWLRISKLRD 317
 Db 202 YGTAQYLLRMLNDHDLAAGRJSTNTVQBTQROHQHGPAGYLEFYEBOG-KMSETLND 260
 Qy 318 TGNALCYTUFVMSL-YLFVITLISYGLMSQLS---EGFGKDGIGLITALMIGLFY 372
 Db 261 LYGLPLITFLMLTILHITVQWMLTILKWTMSLWHALLIAGIATSVNVDLCTMKI 320
 Qy 373 ICDERAHYASVNVRTNFQKULMVELNWMSNDAQTEBIMFLRATENPNSTINCGFFDYNR 432
 Db 321 VGTPARTREESLKT--OKULLRNLSPMDHLKOSIEVPAQTLHQPIETACRMFTLDY 378
 Qy 433 TLFKGLLTMTWVYVLUQFQI 456
 Do 379 TVLFSIAAANTVNLILQPEMAI 402

RESULT 10

Q7PPB5 PRELIMINARY; PRT; 444 AA.
 AC Q7PPB5;
 DT 01-MAR-2004 (TREMBirel. 26, Created)
 DT 01-MAR-2004 (TREMBirel. 26, Last sequence update)
 DB ENSANGP00000004263
 GN Name=ENSGG0000003364;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_TaxID:180454;
 RN [1]
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 DR InterPro:IPR009318; Trichalose_recept.
 DR Pfam:PF06151; Trichalose_recp_1.
 SQ SEQUENCE 444 AA; 51485 MW; 6F09ACD00E059A9 CRC64;

SEQUENCE FROM N.A.

Query Match 5.4%; Score 143; DB 2; Length 444;
 Best Local Similarity 21.5%; Pred. No. 0.023; Gaps 24;
 Matches 98; Conservative 73; Mismatches 170; Indels 114; Gaps 24;

Qy 59 SIGLARLDADEYAPPLD--RKKESTDSTASNNPBPKPSVYRNIDPDPFLRITGVLP 115
 Db 26 STALAE----ETPFDGAERESTHEAVAVIEMGQLE--SLIPIDCAY---- 70

Qy 116 VTHGPARAKAFENNSASPIYS-WVFVFLACVGYVANRTHIVRSLSGPFEAVIAVLP 172
 Db 71 -STDPRDVRMLKRSVOPVYGGVTLFIMLIC---VTAHEPSFGVQOATSVYVA 125

Qy 173 ELV-NLPLMPIMIPLWYEARKIAKLFNDPDEPVLVYQIISHSPLKLRKAWVIAVLP 231
 Db 126 TIVFENFELMLARNN--SQMGWRYTDEAPRPTDPYRPRSTLP--FRKVLHAGFM 181

Qy 232 ILSVSVVITVMSDIA-----INQVPCILDNLNTAMLG 267
 Db 182 PLAFVDTLNPSVASYNLNLIRYCPTAGFMKNFIREHPPVLPVPRVWGMTELT 241

RESULT 11

Q7PKI4 PRELIMINARY; PRT; 416 AA.
 AC Q7PKI4;
 DT 01-MAR-2004 (TREMBirel. 26, Created)
 DT 01-MAR-2004 (TREMBirel. 26, Last annotation update)
 DB ENSANGP00000023329
 GN Name=ENSGG00000020045;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_TaxID:180454;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=PEST;

RA Anopheles Genome Sequencing Consortium;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 DR EMBL: AAB0108960; EAA10739; 2; -
 DR InterPro:IPR009318; Trichalose_recept.
 DR Pfam:PF06151; Trichalose_recp_1.
 SQ SEQUENCE 416 AA; 48071 MW; 1A60309A0F68B1 CRC64;

Query Match 5.4%; Score 142; DB 2; Length 416;
 Best Local Similarity 19.6%; Pred. No. 0.025; Gaps 15;
 Matches 81; Conservative 80; Mismatches 182; Indels 70; Gaps 15;

Qy 100 IDPINWFLR-----IIGVLPV--RKGPARAKFEMNSASFY 134
 Db 12 VSPVWNSVARFESCCMVOVPRWLQVLSQVVKELPFLSPDFR--RERPRNTKDVIL 68

Qy 135 SVVFFVTLACVGYVANRTHIVR-SLGGPFEAVIAVLPNLPIMIPLWTEAKR- 192
 Db 69 STVLLVLYA-IVGPIVCVELXVMSVNOITNALSAQPAFVVFILQVIFRKAC 127

Qy 193 IAKLNDWDPEVLYQIISHSPLKLRKAWVIAVLPILSVISVWVTHVMDLNTQ 252
 Db 128 LHVLLAMEFOLKHVLERMVARM-LEFRYRLFLVKVAT-DILMQLALYTFPDSYTG 185

Qy 253 WVPYCILDNLNTAMLGAWMFLCICBAMSSTAHLLAERFQALKHIGPAM----- 300
 Db 186 RVDYVGLASLTLTXYAMMYPFAVIENFLIAGLIGVAMQMVNGKLRSLRSQSSAAGRQ 245

Qy 301 --VADYRVLMLSKLRTDGNALCYTFVMSLVPPTLISYGLMSOLSEGRGKIDIG 358
 Db 246 PSVLOQYMLHCKNEDMVQKOMETLNFPMLTGMFIMIVSVRPPASV--GNGIAD-- 301

Qy 359 LTITALWNGILF-YICDEAHYASV--INVTNFQKLLKAVLVELNMMNSDAQTE----- 407
 Db 302 -DFKAVINPLPFLFLYQCVQVLLVLLPISVTDHAKMRL-INTVSVNQHHRPQERLV 359

Qy 408 -----INMLRATENPNSTINCGCFPDMVTRLEKGLLTMTWVLUWLUQFOI 454
 Db 360 GREGGYEVJAVDCHQMRNNAINNYYMAMRFLGGMIAATMSVLLQFHI 412

RESULT 12

07PVKO	PRELIMINARY;	PRT;	411 AA.
ID			
AC	07PVK0;		
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
GN	Name=ENSANG60000020468;		
OS	Anopheles gambiae str. PEST.		
OC	Neopatra; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.		
OX	NCBI_TaxID=180454;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PEST;		
RA	Anopheles Genome Sequencing Consortium;		
RU	Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.		
CC	-I - CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary.		
DR	EMBL; AAB0108984; EAA43433.1; -; 8863D72B4E834FBC CRC64;		
SQ	Query Match 5.3%; Score 140; DB 2; Length 382; Best Local Similarity 20.4%; Pred. No. 0.032; Mismatches 155; Indels 84; Gaps 18; Matches 81; Conservative 78; Gaps 18; Best Local Similarity 21.8%; Pred. No. 0.027; Mismatches 87; Indels 87; Gaps 16; Matches 87; Conservative 64; Gaps 16; Best Local Similarity 21.8%; Pred. No. 0.027; Mismatches 161; Indels 87; Gaps 16; Matches 87; Conservative 64; Gaps 16;		
Qy	110 IGVLPVTRHGPRAKAFENNSASPIYSVVFLLACIVGYVANNRTHIVRSLSGPPEAV 160		
Db	17 LVGVPL--RUPRSWEL-----LILCYCTATAGPLVHFLYLAW FVIVM 61		
Qy	169 TAY--IPLVNT-----LPIIMITIPLYW---EARXKULFNDWDDPEVLYQTSGHS- 214		
CC	-I - CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.		
DR	EMBL; AAB0108984; EAA43430.1; -; 6D7A379CB2729ESE CRC64;		
SQ	SEQUENCE 411 AA; 47111 MW; 6D7A379CB2729ESE CRC64;		
Qy	Query Match 5.3%; Score 141.5; DB 2; Length 411; Best Local Similarity 21.8%; Pred. No. 0.027; Mismatches 87; Indels 87; Gaps 16; Matches 87; Conservative 64; Gaps 16;		
Db	46 LGSVNPYFPH-PATLSLPSRSILVYCTIAISPLGRCCVFTTFPMKUGISVVR-YTA 103		
Qy	163 PPEEAIVAYLFVNILPIMI--ILVYEARIAKLUKNDWDDPEVLYQTSGHSLSL-- 217		
Db	104 MIREBLGAIL-----LVPVPCYLCQHAKVQKFE-----LSVEITRSSLPGIF 148		
Qy	218 --KLROKAVY----IAVLPILSVLWSVITVMSDLNINQWVPCILD 260		
Db	149 PFWQWFRTCILCLGSILVSDVSVGVNVAFNLVPLRAGIHKYV-----QIILWSVLS 197		
Qy	261 NIT-----AMIGAWMFLICEAMSITAHLLAERF-----OKALK--HIGPAAVADYRVLWL 309		
Db	198 QITGKSVLILITGTAQVLLMMQDQKHLVAGRLSTMPVQDQHGGAGYLMFYEQ 257		
Qy	307 IWLRLSKLTDPGNALCYTFVFLSMLVPLFTLSTYGM-----SOLSEGGFGKDI 357		
Db	258 LAKVCEFLNDLIGVPLTYFLMLAHLHITFVCULITKLVLRVSYITWQSILLAGIAASY 317		
Qy	358 GITITALWNIGLFYTCDEAHYASVNRTNEQQKKLMLVLMWNSDAQTEINMELRATM 417		
Db	318 VIDLCLLMLRIVGTF----ARTRREESLKT---QKULRLNLSPMDHKLKQSFVPAQTLH 370		
Qy	418 NPSTINCQGFFPVARTIFKGLLITMVYLVLLQFOISI 456		
Db	371 QPIEFTACRMFTLDYTVLFSIAAVTNYLILLIQFENAI 409		

RESULT 13

07PVJB	PRELIMINARY;	PRT;	382 AA.
ID			
AC	07PVJB;		
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	ENSANGP000002448.		
GN	Name=ENSANG60000020468;		
OS	Anopheles gambiae str. PEST.		
OC	Bukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Anopheles.		
OX	NCBI_TaxID=180454;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PEST;		
RA	Anopheles Genome Sequencing Consortium;		
RU	Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.		
CC	-I - CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.		
DR	EMBL; AAB0108986; EAA4331.1; -; 4CF7E1B8DB49A84 CRC64;		
SQ	SEQUENCE 451 AA; 51724 MW; 4CF7E1B8DB49A84 CRC64;		
Qy	Query Match 5.3%; Score 139.5; DB 2; Length 451; Best Local Similarity 18.7%; Pred. No. 0.042; Mismatches 86; Indels 105; Gaps 21; Matches 95; Gaps 21; Best Local Similarity 18.7%; Pred. No. 0.042; Mismatches 174; Indels 105; Gaps 21; Matches 95; Gaps 21;		
Db	72 APPDKDKSSUSTASNKNPFRPSVFRNTIPINWFLRIGVLPVTRHGPRAKAFENNSA 131		
Qy	16 APPSKPGALQATPVGEQEE-OLPHFAFKCFRLFALTPLGLM----DROKDRYVRNTR 69		

This Page Blank (uspto)